

2009 年通信理论与信号处理 学术年会论文集

蔡灿辉 主编

電子工業出版社

Publishing House of Electronics Industry

北京 • BEIJING

内 容 简 介

本论文集收录论文 76 篇, 范围涉及 MIMO 与 OFDM 技术、编码、译码与解调、信道估计与均衡、认知无线电与网络新技术、信息与网络安全、自适应信号处理、阵列信号处理、雷达信号处理、语音信号处理、图像与多媒体信息处理、DSP 应用、基础理论与方法等方面的内容。

本书适合通信、电子、信号与信息处理领域的科研人员 and 高校师生阅读。

未经许可, 不得以任何方式复制或抄袭本书之部分或全部内容。

版权所有, 侵权必究。

图书在版编目(CIP)数据

2009 年通信理论与信号处理学术年会论文集/蔡灿辉主编. —北京: 电子工业出版社, 2009.11

ISBN 978-7-121-09786-7

I .2… II.蔡… III. ①通信理论—学术会议—文集 ②信号处理—学术会议—文集 IV.TN911-53

中国版本图书馆 CIP 数据核字(2009)第 198521 号

责任编辑: 竺南直 特约编辑: 索蓉霞

印 刷: 北京季蜂印刷有限公司

装 订:

出版发行: 电子工业出版社

北京市海淀区万寿路 173 信箱 邮编 100036

开 本: 787×1 092 1/16 印张: 32.5 字数: 832 千字

印 次: 2009 年 11 月第 1 次印刷

定 价: 108.00 元

凡所购买电子工业出版社图书有缺损问题, 请向购买书店调换。若书店售缺, 请与本社发行部联系, 联系及邮购电话: (010) 88254888。

质量投诉请发邮件至 zlt@phei.com.cn, 盗版侵权举报请发邮件至 dbqq@phei.com.cn。

服务热线: (010) 88258888。

中国通信学会通信理论与信号处理专业委员会
2009 年通信理论与信号处理学术年会

主办 (Sponsor):

中国通信学会通信理论与信号处理专业委员会

承办 (Organizer):

华侨大学

大会主席:

张乃通 (教授 中国工程院院士)

丘 进 (教授 华侨大学校长)

郑宝玉 (教授 中国通信学会通信理论与信号处理专业委员会主任)

程序委员会

主 席: 郑宝玉 (教授 中国通信学会通信理论与信号处理专业委员会主任)

副主席:

项海格 朱光喜 林孝康 吴镇扬 彭启琮 阮秋琦 范平志 田宝玉

委员 (按姓氏拼音顺序排列):

安建平 戴旭初 杜 歆 范 京 冯穗力 葛临东 顾学迈 何 晨
侯春萍 胡 波 黄剑明 刘贵忠 刘文予 刘 中 仇洪冰 宋志群
孙 洪 王 晖 王盛利 王晓明 王岩飞 谢显中 许录平 许宗泽
颜永红 殷福亮 于 全 张邦宁 张晓琳 郑林华 周治中 竺南直
崔景伍

组织委员会

主 席: 蔡灿辉 (教授 华侨大学信息科学与工程学院)

副主席: 苏凯雄 (教授 福州大学) 王琳 (教授 厦门大学)

委 员: 陈忠辉 程恩 冯桂 黄联芬 凌朝东 许力 崔景伍 陈东华 陈燕

前 言

当前，网络和通信的发展趋势是宽带化、移动化、泛在化，人们将通过 RFID（射频识别芯片）和传感器生活在无所不在（Ubiquitous）的网络（即所谓泛在网络或 U 网络）中。移动化催生 U 网络时代；融合与协同，是 U 网络架构的主要特征。频谱利用率是考验新一代移动通信网络的重要指标。为了满足新一代移动通信系统对频谱效率的要求，人们纷纷探索更加高效的无线通信技术。认知无线电技术是有效提高频谱利用率的重要手段，是未来无线移动通信的发展方向之一；已经提出的高效传输技术有多载波调制技术、多天线 MIMO 技术、空时信号处理技术、混合多址技术等。它们都是信号处理和无线通信（包括组网技术）的结合，这些结合极大地推动了无线通信技术的进步。

所有这些在本论文集集中均有所反映。此外，论文集集中还包括信道估计、均衡与编码，计算智能与盲信号处理，雷达信号处理与智能天线，信息论、网络编码及安全，DSP 及应用，语音与图像信号处理，光通信与光信息处理等。共 76 篇。这些论文是通过专家严格评审，从收到的 149 篇挑选出来的。这些论文充分反映了通信与信号处理领域近年来的研究成果及今后的发展趋势。由于版面和时间等条件的限制，还有不少优秀论文未能选入，深表歉意。

本着通信理论与信号处理专业委员会的一贯宗旨，我们希望本次学术会议也像本专委会的历次学术活动那样，办成一次既是交流成果，联络感情，结识新友的会议，也是探索未来，促进通信与信号处理学科发展的会议。衷心祝愿大家在会议期间收益多多，满载收获地返回各自地工作岗位，创造新的业绩，做出新的贡献。

最后，感谢大家来参加这次盛会，感谢东道主和会议的工作人员为大会的成功召开所付出的巨大努力，感谢电子工业出版社各级领导和编辑为本论文集的出版所付出的辛勤劳动。

郑宝玉

中国通信学会通信理论与信号处理专业委员会主任委员

目 录

一种基于 Alpha 稳定分布的大气噪声生成算法及应用	马金全 葛临东 童 莉 (1)
SAR 方位预处理的两级滤波实现方法	王岩飞 刘 畅 (8)
一种提高 GNSS 卫星信号模拟器伪距精度的方法	彭 鸣 吴晓非 邹永忠 (13)
相干信源的二维 DOA 盲估计算法	余 俊 陈未央 张小飞 (19)
一种新的低功率信号检测算法	王 磊 郑宝玉 崔景伍 (25)
宽带雷达信号接收波束形成的方法	章华奎 王盛利 (33)
基于瞬时距离-瞬时多普勒的高速机动目标成像	句彦伟 李士国 于 立 (41)
基于 Markov 随机场和模拟退火算法的 SAR 图像噪声抑制	钱 博 (47)
循环相关谱检测方法及其门限的确定	赵知劲 张伟卫 徐世宇 (54)
GPS 信号捕获中信噪比提高的研究	雷 蕾 林静然 彭启琮 (61)
基于量子遗传神经网络的 MIMO 信号检测技术	周 敏 李 飞 (67)
基于最大似然聚类的 GMM 优化方法及其在说话人辨认中的应用	胡 婕 周 琳 (73)
一种新的基于段长分布的语音识别模型	杜树木 何良华 (80)
基于内容分析和点击率记录的混合音乐推荐系统	张 燕 李燕萍 (85)
能量重心校正法估计 MPSK 信号载波频差	张冬玲 葛临东 (93)
基于光流场法的目标跟踪算法在 DM642 上的高效实现	武 曜 管 庆 徐 胜 (98)
目标跟踪算法在 DM6437 上的实现和优化	赵 渊 管 庆 徐 胜 (107)
一种新型格型 IIR 滤波器的研究与设计	傅文渊 凌朝东 (114)
多视点图像的本征图分解	施柏鑫 李扬曦 许 超 (119)
基于行移字移编码算法的文本水印鲁棒性研究	黄兴莉 邵怀宗 彭启琮 (124)
随机密钥预分布的抗攻击模型研究	张 啸 王 晖 林晓辉 (130)
基于加权二乘向量机的图像增强算法	彭代强 林幼权 (141)
基于置信度传播的 MSTAR 图像的分割算法	徐 戈 黄志强 (147)
一种适合子带 LMS 的非均匀滤波器组	张 炜 许小东 戴旭初 (154)
3780 点 FFT 的 FPGA 设计与实现	陈联武 苏凯雄 (162)

一种基于 D-S 证据理论的高光谱图像分类方法	余二永	(167)
简化的 SMV 算法及其仿真	王小兵 黄 冰 杜俊宇	(176)
基于方向特征和局部 walsh 变换的虹膜识别新算法	王凤随 李素文 冯 桂	(182)
变长小生境算法用于图像超分辨率复原	张煜东 吴乐南 奚 吉 王水花	(189)
一种基于形状信息的前景检测方法	刘乐元 桑 农	(200)
多模态导航定位中星座选择算法	邵怀宗 邵 龙 彭启琮	(207)
H.264/AVC 的帧内模式选择优化算法	李仲林 胡 栋	(212)
基于 CORDIC 算法的数字下变频器实现	王小凡 唐友喜	(218)
对角空时码在无线中继网络中的应用	赵知劲 杨进丽 王海泉	(223)
认知无线网络中一种自适应 TDD 传输策略的研究	陈晓辉 孟庆民 郑宝玉	(228)
一种基于延时的 AODV 路由协议的改进	陈哲愚 张 建 陈 燕	(234)
一种改进的 BP 量子 LDPC 译码算法	段 灿 朱修利	(240)
基于机会式网络编码的两用户协作	李世唐 郑宝玉	(245)
一种多无线电多信道环境下新的信道状态预测算法	黄 川 郑宝玉	(251)
协作通信的能量有效性和时间有效性研究	季 薇 郑宝玉 江若宜	(257)
时间选择性衰落信道的基扩展模型比较	陈东华 蔡灿辉	(265)
多小区 OFDMA 系统下行链路的分布式子载波分配	邱 涛 贺志强 牛 凯 田宝玉	(270)
一种下行单用户协作多点传输的调度算法	张成强 贺志强 牛 凯 田宝玉	(278)
一种基于图码的量子稳定子码构造方法的研究	朱修利 赵生妹	(283)
基于 SFBC-OFDM 系统的时变信道估计和 信号检测联合算法	任术波 郭俊奇 项海格	(288)
基于预测的 OFDMA 系统中断容量的优化	吴 波 沈 骏 陈 晨 项海格	(297)
一种基于预滤波的 NC-OFDM 信号同步参数盲估计算法	吴佩霖 许小东 戴旭初	(305)
认知无线网络频谱共享的跨层设计	张伟青 叶 梧 冯穗力	(312)
多射频多信道无线 Mesh 网络的跨层公平资源分配	黄 鑫 冯穗力 庄宏成	(319)
DF 协作中继网络基于最小能量成本的中继选择和功率控制策略	柯 峰 冯穗力	(328)
在 WSNs 中基于覆盖与连通的分簇路由协议的改进	朱丽娟 王敬国 林晓辉 王 晖	(334)

基于量子遗传算法的 OFDM 自适应调制技术.....	李兆华	李 飞	(342)
基于蚁群算法的 OFDMA 系统自适应资源分配算法			
.....	张 昕	叶 梧	冯穗力 庄宏成 (348)
用于多小区 OFDMA 系统的基于粒子群的分布式功率分配方法			
.....	孙文文	贺志强	牛 凯 田宝玉 (356)
一种适合于低压 PLC 的组合译码 HARQ 算法	朱子善	林 东	(364)
一种基于信道可靠性的协作频谱感知算法	岳文静	郑宝玉	(370)
认知无线网络中基于博弈论的功率控制算法	王正强	蒋铃鸽	何 晨 (378)
多层网状局域连通无线传感器网络定位分析	刘 枫	郭爱煌	杭海存 (385)
DVB-RCS 中循环递归 Turbo 码尾比特译码技术.....	巩克现	刘世刚	葛临东 (391)
基于协同中继的异构无线网络切换性能分析			高智伟 (398)
分簇无线传感器网络中最小退避窗口自适应机制	林 颖	许 力	林 晖 (405)
一种解决认知无线电频谱分配的量子遗传算法	郭 明	李 飞	(410)
一种求解最短路径路由问题的量子遗传算法	刘 欣	李 飞	张 杰 (416)
单载波块传输系统中稀疏信道估计算法研究	黄 帅	朱卫平	孟庆民 (422)
基于无扰动的最大速率传输 MIMO 系统半盲信道估计.....	陈 钟	朱卫平	(427)
数字集群系统中移动链路管理的研究与实现	张 建	何志敏	(433)
数字无中心移动通信系统 MAC 层的协议设计			强文炜 (440)
基于 802.16e 交织编码部分的 FPGA 设计与实现	毛智强	雷维嘉	谢显中 (447)
基于 DTMB 标准的卷积交织仿真与实现.....	肖剑辉	叶宇煌	(455)
无线传感器网络中基于网络编码的信道 自适应协同传输方案.....	张 良	蔡跃明	(462)
双向中继对秩亏 MIMO 信道容量的影响.....	屠 佳	蔡跃明	(469)
基于 ENC28J60 芯片的以太网接口方案设计.....	付衍荣	房 颖	(476)
基于多频带 OFDM 的矿井巷道超宽带无线通信建模			
.....	钱建生	王 鑫	孙彦景 李世银 (482)
序列相关性对系统性能影响的理论分析和评估	吴 娜	吴晓非	邹永忠 (490)
基于中继节点的无线传感器网络分簇路由算法	李 雷	付东阳	李艳辉 (497)
DSTFT 解调 FSK 信号中窗函数的影响.....	陈大平	陈忠辉	(503)

一种基于Alpha稳定分布的大气噪声生成算法及应用

马金全 葛临东 童 莉

(信息工程大学 信息工程学院, 河南 郑州, 450002)

摘 要: 大气噪声是短波天波信道的主要影响因素之一, 如何生成符合实际信道特性的大气噪声对于宽带短波信道模拟器研制具有重要作用。在分析 α 稳定分布定义与性质的基础上, 针对大气噪声特性, 提出了一种基于 α 稳定分布的大气噪声生成算法, 并在宽带短波信道模拟器中进行了应用。实验结果表明, 利用该方法产生噪声不仅更加符合信道实际情况, 而且可以有效降低模拟器的实现复杂度。

关键词: 通信技术; 信道模拟; Alpha 稳定分布; 宽带短波信道; 大气噪声; 软件无线电

An Algorithm of Atmospheric Noise Generating Based on Alpha Stable Distribution

Ma Jin quan Ge Lin dong Tong Li

(University of Information Engineering, Zhengzhou, 450002, China)

Abstract: As atmospheric noise being one of the main factors influencing HF channel characteristics, the method of generating atmospheric noise in line with real channel plays an important role in the research of wide-band HF channel simulation system. By introducing the alpha stable distribution and aiming at the atmospheric noise property, a noise generating algorithm based on alpha stable distribution is suggested and applied in the wide-band HF channel simulation system. The experiment results show that the algorithm could not only generate atmospheric noise more conforming to reality, but predigest the simulation system with better performance.

Keywords: communications technology; channel simulation; Alpha stable distribution; wide-band HF channel; atmospheric noise; software radio

引言

在短波通信设备研制开发过程中, 利用信道模拟器进行通信仿真是一种重要手段。随着短波通信的宽带化, 研制宽带短波信道模拟器日显重要。

目前, 国内已实现带宽 3kHz 的短波信道模拟器^[1], 国外 John F. Mastrangelo 等人根据 Vogler 短波信道模型开发了带宽 1MHz 的短波信道模拟器^[2], 基本达到短波通信设备开发中的测试要求。但由于利用天波传输的短波通信受到电离层影响很大^[3], 这些方案产生的循环噪声与实际大气噪声差别较大, 不能有效模拟实际情况。

本文提出一种基于 α 稳定分布的大气噪声计算方法, 所有噪声信号都根据相应的参数实

时计算产生，能提高宽带模拟器的工作效率，降低实现复杂度，利于用户操作。

1 Vogler信道模型及噪声特点

Vogler 短波信道模型又称为 ITS 模型，是在 Vogler, Hoffmeyer 等人研究成果基础之上提出的一种短波信道参考模型^[3]。该模型数学表达为

$$y(t) = x(t) * h(t, \tau) + n(t) \quad (1)$$

式中 $h(t, \tau)$ 是信道冲击响应； $n(t)$ 表示干扰及噪声模型，由大气噪声、带限高斯噪声、单音干扰及人为脉冲干扰几部分组成，其等效低通形式为

$$\begin{aligned} v(t) = & a(t) + g(t) + \\ & \sum_i A_i \exp[j(2\pi\Delta f_i t + \phi_i)] + \\ & \sum_k B_k \frac{\sin[2\pi W(t - t_k)]}{t - t_k} \exp(j2\pi f_0 t_k) \end{aligned} \quad (2)$$

式(2)中 $a(t)$ 表示大气噪声； $g(t)$ 表示零均值的带限复高斯噪声；第三部分表示窄带干扰；第四部分表示人为脉冲干扰。以往的信道模拟中，一般假设 $v(t)$ 是大量统计独立且有限方差随机变量的叠加组合，因而 $v(t)$ 服从高斯分布。

实际中，大气噪声主要由雷电瞬时放电引起，其分布具有较厚的拖尾，是一种具有较大波动的高可变的过过程，往往不服从高斯分布。特别是在无线电干扰较少的地区或频率高于 10MHz 的情况下，大气噪声对于信道的影响非常明显^[4]。若只采用高斯分布模型来描述信道噪声，将会因模型与信号噪声不能很好匹配而导致所设计信号处理器性能退化，这对于通信设备研制是不利的。

广义中心极限定理表明^[5]：对于独立同分布随机变量的和，无论各个随机变量是否存在有限方差，当变量数目无限增加时，必将收敛于 α 稳定分布族。Nikias 等人的研究表明，利用 α 稳定分布模型方可更加有效地描述具有尖峰脉冲特性的噪声。

2 α 稳定分布理论

2.1 α 稳定 (αS) 分布的定义^[5]

α 稳定分布一般由其特征函数给出。如果存在参数 $0 < \alpha \leq 2, \sigma > 0, -1 \leq \beta \leq 1$ 和实数 μ ，使得随机变量 X 满足如下的特征函数，称 X 服从稳定分布，记 $X \square S_\alpha(\sigma, \beta, \mu)$ ：

$$E[e^{i\theta X}] = \begin{cases} \exp\left\{-\sigma^\alpha |\theta|^\alpha \left[1 - i\beta \tan \frac{\alpha\pi}{2} \cdot \text{sgn}(\theta)\right] + i\mu\theta\right\} & (\alpha \neq 1) \\ \exp\left\{-\sigma |\theta| \left[1 - i\beta \frac{2}{\pi} \ln|\theta| \cdot \text{sgn}(\theta)\right] + i\mu\theta\right\} & (\alpha = 1) \end{cases} \quad (3)$$

其中 $\text{sgn}(\theta)$ 是符号函数。

α 稳定分布四个参数具有明确的意义：

- 特征因子 α ，又称为特征指数，决定分布脉冲特性程度。 α 越小，所对应的分布具有较重的拖尾；当 $\alpha \leq 1$ ，分布具有无限的均值和方差。
- 尺度参数 σ ，又称为分散系数，是样本相对于均值分散程度的度量，类似于高斯分布的方差；当 $\alpha = 2$ 时， σ 取值是方差值的一半。
- 偏斜参数 β ，确定分布的斜度。
- 位置参数 μ ，对于 $S\alpha S$ 分布， $0 < \alpha \leq 1$ 时， μ 表示中值；当 $1 < \alpha \leq 2$ 时， μ 表示均值。

稳定分布又可分为对称分布 ($S\alpha S$) 和非对称分布两大类。当 $\beta = 0$ 时， $S_\alpha(\sigma, \beta, \mu)$ 关于 μ 是 $S\alpha S$ 对称的。

特殊地，当 $\alpha = 2$ 时，随机变量 X 服从均值为 μ 方差为 $2\sigma^2$ 的高斯分布；当 $\alpha = 1$ 且 $\beta = 0$ 时， X 服从柯西分布。

2.2 α 稳定分布的不同参数系

对 α 稳定分布的描述存在着不同的参数系，就随机变量的产生而言，重要的有三种参数系^[6]：由式(3)定义给出的标准参数系 (S)，Nolan 提出的参数系 (S^0) 和 Zolotarev 的参数系 (S^2)。

S^2 参数系提出的主要目的是利于理论分析与推导，其特征函数有如下形式^[7]：

$$E[e^{i\theta X}] = \begin{cases} \exp\left\{-\sigma_2^\alpha |\theta|^\alpha \exp[-i\beta_2 \operatorname{sgn}(\theta) \frac{\pi}{2} K(\alpha)] + i\mu\theta\right\} & (\alpha \neq 1) \\ \exp\left\{-\sigma_2 |\theta| \left[\frac{\pi}{2} + i\beta_2 \ln|\theta| \operatorname{sgn}(\theta)\right] + i\mu\theta\right\} & (\alpha = 1) \end{cases} \quad (4)$$

其中

$$K(\alpha) = \alpha - 1 + \operatorname{sgn}(1 - \alpha) = \begin{cases} \alpha, & \alpha < 1 \\ \alpha - 2, & \alpha > 1 \end{cases}$$

S^2 参数系的 α 和 μ 同标准参数系所定义参数相同，而 σ_2 和 β_2 同 σ 和 β 的关系为：

$$\sigma_2 = \begin{cases} \sigma \left(1 + \beta^2 \tan^2\left(\frac{\pi\alpha}{2}\right)\right), & \alpha \neq 1 \\ \frac{2}{\pi} \sigma, & \alpha = 1 \end{cases} \quad (5)$$

$$\beta_2 = \begin{cases} \frac{2 \arctan[\beta \tan(\frac{\pi\alpha}{2})]}{\pi K(\alpha)}, & \alpha \neq 1 \\ \beta, & \alpha = 1 \end{cases} \quad (6)$$

S^0 参数系表征的稳定分布的特征函数、概率密度函数和累积分布函数关于所有的四个参数均连续，其参数 α 、 σ 和 β 与标准参数系相同。文献[6]证明， S^0 参数系和 S^2 参数系均同标准参数系存在相应的变换关系。

2.3 α 稳定分布的性质

$S_\alpha(\sigma, \beta, \mu)$ 稳定分布具有的性质较多，这里仅给出 3 个基本性质：

性质 1: 若 $X \square S_\alpha(\sigma, \beta, \mu)$, c 是一个实常数, 则 $X + c \square S_\alpha(\sigma, \beta, \mu + c)$

性质 2: 若 $X \square S_\alpha(\sigma, \beta, \mu)$, b 是一个非零的实常数, 则

$$bX \square \begin{cases} S_\alpha(|b|\sigma, \text{sgn}(b)\beta, b\mu), & \alpha \neq 1 \\ S_1(|b|\sigma, \text{sgn}(b)\beta, b\mu - \frac{2b}{\pi} \ln|b|\sigma\beta), & \alpha = 1 \end{cases}$$

性质 3: 令 X_1 和 X_2 为独立的随机变量, 且满足 $X_i \square S_\alpha(\sigma_i, \beta_i, \mu_i), i=1,2$, 则随机变量 $X_1 + X_2 = X \square S_\alpha(\sigma, \beta, \mu)$, 其中

$$\sigma = (\sigma_1^\alpha + \sigma_2^\alpha)^{1/\alpha}, \beta = \frac{\beta_1 \sigma_1^\alpha + \beta_2 \sigma_2^\alpha}{\sigma_1^\alpha + \sigma_2^\alpha}, \mu = \mu_1 + \mu_2$$

3 α 稳定分布随机变量的产生

3.1 服从 αS 分布随机变量的产生

仿真生成服从参数 $0 < \alpha \leq 2, \sigma > 0, -1 \leq \beta \leq 1$ 和实数 μ 的稳定分布关键在于产生服从 $S_\alpha(1, \beta, 0)$ 的随机变量。通常采用 Zolotarev 定理产生 $S_\alpha^2(1, \beta_2, 0)$ 分布, 再根据式(5)和式(6)生成 $S_\alpha(\sigma, \beta, \mu)$ 。Zolotarev 定理如下:

若 γ 是 $(-\pi/2, \pi/2)$ 上的均匀分布, W 是一个均值为 1 的服从指数分布的随机变量, 并且 W 和 γ 相互独立, 则 $\alpha \neq 1$ 时, 随机变量 X 服从 $S_\alpha^2(1, \beta_2, 0)$ 分布, 当且仅当

$$X = \frac{\sin(\gamma - \gamma_0)}{(\cos \gamma)^{1/\alpha}} \left\{ \frac{\cos[\gamma - \alpha(\gamma - \gamma_0)]}{W} \right\}^{\frac{1-\alpha}{\alpha}} \quad (7)$$

$\alpha = 1$ 时, 随机变量 X 服从 $S_1^2(1, \beta_2, 0)$ 分布, 当且仅当

$$X = \left(\frac{\pi}{2} + \beta_2 \gamma \right) \tan \gamma - \beta_2 \log \left(\frac{W \cos \gamma}{\frac{\pi}{2} + \beta_2 \gamma} \right) \quad (8)$$

其中 $\gamma_0 = -\frac{\pi}{2} \beta \frac{K(\alpha)}{\alpha}$ 。

结合性质 1、2 和式(5)、式(6), 可有标准参数系下 α 稳定分布仿真公式: 若 $X \square S_\alpha^2(1, \beta_2, 0)$ 则

$$U = \sigma_2 X + \mu \square \begin{cases} S_\alpha(\sigma, \beta, \mu), & \alpha \neq 1 \\ S_\alpha(\sigma, \beta, \mu - \frac{2}{\pi} \sigma \beta \ln(\frac{2}{\pi} \sigma)), & \alpha = 1 \end{cases} \quad (9)$$

仿真时先产生 $X \square S_\alpha^2(1, \beta_2, 0)$, 再由式(9)产生服从标准参数系下任意参数取值的 α 稳定分布的随机变量 U (包括服从 $S\alpha S$ 分布的情况)。

3.2 MATLAB 仿真与分析

基于上述生成算法, 利用 MATLAB 进行 1000 次 Monto-Carlo 仿真, 产生长度为 5000 的 α 稳定分布的随机序列。

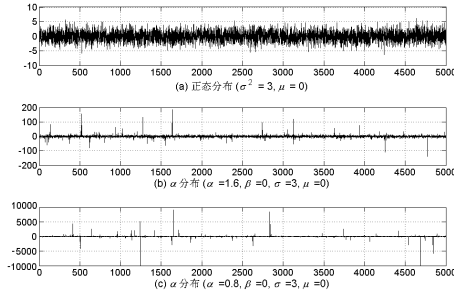


图 1 不同情况下的噪声序列

图 1 为三种情况下产生的噪声，其中(a)为方差 $\sigma^2 = 3$ ，均值 $\mu = 0$ 的高斯噪声序列，(b)为 $\alpha = 1.6, \beta = 0, \sigma = 3, \mu = 0$ 的 α 分布噪声序列，(c)为的 $\alpha = 0.8, \beta = 0, \sigma = 3, \mu = 0$ 的 α 分布噪声序列。可以看出， α 分布噪声序列具有明显的非高斯脉冲特性，而且随着 α 的减小，尖峰脉冲特性愈加明显。

图 2 为不同 β 值和 σ 值时的 $S_\alpha(\sigma, \beta, \mu)$ 和 $S\alpha S$ 分布随机变量的概率密度函数（PDF）估计的对比图。可以看出，仿真产生的随机变量的 PDF 与理论分析结果一致。

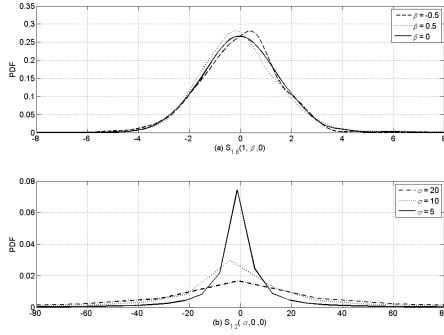


图 2 不同 β 与 σ 值下随机变量的 PDF 估计曲线

3.3 信噪比的确定

由于 α 稳定分布不存在有限的二阶矩，致使噪声的方差变得没有意义，因此通常采用混合信噪比设定信号与噪声的功率比：

$$\text{MSNR}_{dB} = 10 \log_{10} \left(\frac{\sigma_s^2}{\sigma_\alpha} \right) \quad (10)$$

式中， σ_s^2 和 σ_α 分别表示高斯信号的方差和 α 稳定分布噪声的尺度参数。假定要对给定信号 $s(n)$ 和 α 稳定分布噪声 $v(n)$ 设定混合信噪比为 $\text{MSNR}_{dB} = m\text{dB}$ ，则按照此信噪比大气噪声应满足

$$\sigma_\alpha = 10^{-m/10} \sigma_s^2 \quad (11)$$

4 在宽带短波信道模拟器中的应用

4.1 宽带短波信道模拟器结构分析

John F.Mastrangelo 等人采用了 VME 总线加存储单元的方式，结构较为复杂。本文的宽带短波信道模拟器采用 cPCI 总线架构，结构更加简洁灵活，其功能结构图如图 3 所示。

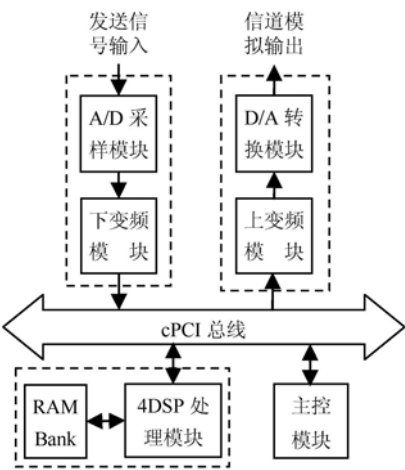


图 3 宽带短波信道模拟器功能结构图

信号在 A/D 采样后，经下变频转为基带信号，经由 cPCI 总线传至信道模拟模块；模拟模块核心为 4DSP 处理单元（ADSP TS201S），可实现信道模拟、噪声叠加以及信噪比计算等功能，噪声数据在系统启动后由主控模块转存到 256M RAM 空间；主控模块负责系统参数设置、数据保存和软件界面显示等功能；信号与噪声的合成数据再经 cPCI 总线传至输出模块，最终形成信道模拟信号。

4.2 大气噪声的生成

实时生成大气噪声数据，需要进行大量繁杂的指数及乘法运算，为减小计算量，大气噪声数据事先由 MATLAB 产生，以文件形式存储在 RAM 块中。为节省空间，RAM 块中存储了 2 个 1M 数据块 A 和 B，每次分别按照均匀分布随机地从数据块 A 和数据块 B 的 1M 数据中取出 1 个数据 X_A 和 X_B ，根据 α 稳定分布的性质 3，由 $aX_A + bX_B$ （ a, b 为常数）形成的噪声数据仍服从 α 稳定分布，故由 2 个 1M 数据可产生 10^{12} 的噪声数据（在 20MSPS 采样率下可持续约 13 小时）。因此只要存储的噪声数据足够多，就可保证在长时间内噪声数据不重复，不仅使噪声更加接近实际情况，还可以避免大量的数学运算，提高模拟器工作效率。

模拟信道时，均匀分布随机变量采用线性同余法产生，其周期与噪声数据长度一致；高斯噪声、单音干扰及人为脉冲干扰采用文献[1]的方法产生。

图 4 为信噪比 5dB 时模拟器产生衰落信号与不同噪声的叠加输出（采样率 20MSPS，样点数为 5000）。其中(a)为信号叠加高斯噪声，(b)为信号叠加大气噪声（ $\alpha=1.8$ ），(c)为信号叠加高斯噪声与大气噪声（高斯噪声与大气噪声功率比为 20dB）。

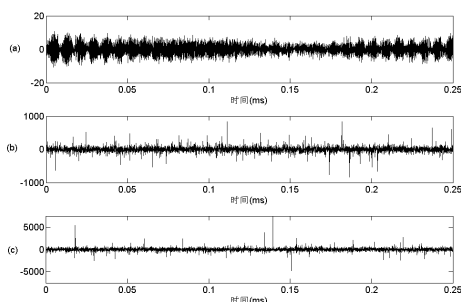


图 4 SNR=5dB 时模拟器不同的输出信号

从图中可以看出,信道模拟器加入大气噪声后,即使大气噪声不在总噪声中占据主导地位,但是其尖峰脉冲特性对信道的影响是显著的,在信道模拟器中加入大气噪声更加符合实际情况。

5 结论

本文提出了一种基于 α 稳定分布的大气噪声生成算法,该方法可以使模拟的信道噪声更加接近实际情况,也能简化宽带短波信道模拟器结构及操作。在宽带短波信道模拟器中,噪声信号计算采用了该方法,解决了实时计算的问题,并且简化了结构,达到了预期的效果。

参 考 文 献

- [1] 马金全,杜栓义,邱长兴. 基于 DSP 的话音带宽短波信道模拟器[J]. 电子技术应用, 2005,31(3):56-58.
- [2] Mastrangelo J F. A new wideband high frequency channel simulation system[J]. IEEE Transactions on Communications, 1997,45(1):26-34.
- [3] Vogler L E, Hoffmeyer J A. A model for wideband HF propagation channels[J]. Radio Science, 1993,28(6):1131-1142.
- [4] 胡中豫. 现代短波通信[M]. 北京:国防工业出版社, 2005.
- [5] 邱天爽,张旭秀等. 统计信号处理——非高斯信号处理及其应用[M]. 北京:电子工业出版社, 2004.
- [6] 李旭涛,朱光喜等. Alpha 稳定分布的参数表征及仿真[J]. 信号处理, 2007,23(6):814-817.
- [7] V. M. Zolotarev. One-dimensional Stable Distributions[M], Amer. Math. Soc. Transl. of Math. Monographs, Vol. 65. Amer. Math. Soc., Providence, R. I.(Transl. of the original 1983 Russian).

作者简介:

马金全 (1975.3-), 男, 甘肃张掖人, 博士研究生, 信息工程大学信息工程学院讲师, 主要从事软件无线电、无线通信信号处理研究。

葛临东 (1946.8-), 男, 博士生导师, 信息工程大学信息工程学院教授, 主要从事软件无线电、无线通信信号处理研究。

童莉 (1978.11-), 女, 湖北荆州人, 博士, 信息工程大学信息工程学院讲师, 主要从事信号与信息处理研究。

SAR方位预处理的两级滤波实现方法

王岩飞 刘 畅

(中国科学院电子学研究所 北京海淀区北四环西路 19 号, 100190)

摘 要: 在合成孔径雷达实时成像处理器中, 通常采用方位向预滤波降采样技术, 在保证信号性能的同时减少数据量以降低对成像处理器运算和存储等能力的要求。针对实际系统的实现要求, 论文提出了一种两级滤波降采样的实现方法。通过分级处理, 在简化滤波运算的同时降低了初始阶段的雷达数据率, 并通过后续的处理来保证雷达的成像质量。给出的方法适合实时 SAR 成像处理的应用, 实际数据的处理结果验证了方法的有效性。

关键词: 信号处理; 合成孔径雷达; 预滤波; 实时成像处理

A Two-step Azimuth Prefilter Method for Real-time SAR Imaging Processor

Wang Yan fei Liu Chang

Institute of Electronics, Chinese Academy of Sciences, Beijing 100190, China

Abstract: In SAR signal processing, the azimuth prefiltering is usually used to reduce the computation and data storage requirement without influence on the effective Doppler band. This paper provides a new two-step filtering method to achieve better processing performance. During the first step, data rate is cut to half with short filtering, and a second suitable filter is further applied to keep the imaging resolution. The proposed prefilter algorithm is demonstrated with real SAR data, and is especially valuable for SAR real time processor.

Keywords: signal processing; SAR; prefilter; real-time processor

1 引言

合成孔径雷达(SAR)是一种对地面目标进行成像观测的微波遥感手段。具有能够全天候工作的优点, 因而广泛应用于国防、自然资源勘探、灾害监测等领域。实时成像处理器通过对 SAR 系统采集到的目标回波信号进行处理实时形成雷达图像, 对于具有时效性要求的应用而言是 SAR 系统必不可少的重要组成部分。在通常的机载 SAR 系统中, 为了降低方位模糊, 提高图像的信噪比, 雷达发射的探测脉冲的频率往往高于要求的脉冲频率, 等效于方位向系统的带宽远高于信号的带宽。考虑到 SAR 的数据量比较大, 并且有效信号的带宽相对较低, 因此在成像处理器中通常采用方位向预滤波降采样技术来减少数据量, 降低对处理器运算、存储等能力的压力, 同时保证雷达图像的性能指标^{[1][2]}。

SAR 实时成像处理器的方位预滤波通常采用 FIR 滤波器, 考虑到运算和缓存量的限制,

滤波器阶数的选择一般都会比较适中^[3]。一方面,阶数不能过大,以免增加过多的计算量和数据缓存量;另一方面,还要在有限的阶数条件下,设计尽可能优化的滤波器。本文基于实际系统的实现要求,提出了一种两级滤波降采样的实现方法。首先对雷达方位向信号进行低阶数的滤波处理,由于阶数较低,计算量和数据缓存量都相对较小,滤波器的性能也由此下降,例如,滤波器的过渡带变宽等,但有效信号的频谱性能可以得到保证。在第一级滤波降采样后,可以减少至少一倍的数据量,降低对处理器的处理运算、缓存等要求的压力。尽管降采样可以带来频谱混叠等影响,但同样要求保证有效信号的性能。之后进行第二级高性能滤波,提取有效的信号,经过后续处理,可以得到质量有保证的雷达图像。

本文的以下部分,首先介绍两级滤波降采样方法的基本思路,之后结合实际系统给出一个两级滤波降采样的设计,并利用实际数据进行处理验证,最后对本文两级滤波降采样的方法进行总结和讨论。

2 两级滤波降采样方法

SAR 系统是一种对地面目标进行二维成像的微波雷达,雷达通常装载于飞机等运动平台上,利用脉冲压缩技术获得距离向的高分辨率,利用孔径合成技术获得方位向的高分辨率,其中沿飞机飞行的方向称为方位向,垂直于飞机航线的方向称为距离向。典型的 RD 成像算法是将接收到的雷达信号经 A/D 采样后,首先进行距离压缩,之后进行方位向预滤波降采样,然后进行转置存储,再进行方位向压缩,经过逆转置存储后将图像连续地进行显示^[4]。

方位预处理一般由低通滤波器和重抽样两部分组成。滤波器主要用于去掉无用的信号,避免在降采样时带来信号混叠的影响;重抽样过程则是通过降低采样率的过程来减少数据量。假设方位向的采样间隔为 Δ_x ,在距离为 R 的目标的第 n 次回波用 $r(n\Delta_x)$ 来表示,则预处理滤波器的输出为

$$q(k\Delta_x) = \sum_{m=0}^{N-1} r(k\Delta_x - m\Delta_x)y(m\Delta_x) \quad (1)$$

式中 $y(m\Delta_x)$ 为滤波器的系数。滤波后的信号再以较低的采样频率降采样得到 $q(n\gamma\Delta_x)$,其中 γ 为降采样系数。

由于 SAR 的回波数据是随着发射脉冲一条距离线接着一条距离线顺序地到达,所以要进行方位向的滤波处理必须要对数据进行临时存储。假设方位向的滤波器阶数为 N ,通常情况下需要存储 N 条距离线的数据,并进行相应的滤波运算。可见滤波阶数越大,对预处理提出的存储和运算要求也越高。为了减轻对处理器的要求,通常希望滤波器阶数尽可能少。然而,滤波器的性能通常是与滤波器的阶数直接相关的^[5]。加窗函数可以提高滤波器的性能。但是,在提高峰值旁瓣比和降低阻带衰减的同时,也会带来主瓣展宽和过渡带变宽的问题。

针对上述问题,本文提出了采用两级滤波降采样的方法。在第一级滤波过程中采用阶数比较低的滤波器,在保证有用信号性能的前提下,尽可能减少对处理器的运算、存储等性能要求。如图 1 和图 2 所示,图中纵坐标为窗函数的频域幅度谱的分贝表示,横坐标为归一化的频率表示,图 1(a)为 55 阶海明滤波器,图 2(b)为 9 阶海明滤波器。图 2 为对应图 1 滤波器的 1/2 重抽样后的频域幅度谱。从图 1 和 2 中可以看出,与通常的理论分析一样,阶数高的滤波器相对于阶数低的滤波器具有更好的性能;经过抽样以后,阶数高的滤波器混叠影响比较

小，阶数低的滤波器混叠影响比较大。然而，分析图 2(a)和图 2(b)还可以发现，如果有效信号的频带宽度在 0.1 以内时，则尽管滤波器的阶数比较低，但图 2(b)中混叠带来的对有用信号的影响也比较小。如果在后续的滤波处理中，例如进行第二次滤波或者结合方位压缩处理等，能够有效地滤除有效信号之外的无用信号，则同样可以获得比较好的效果。

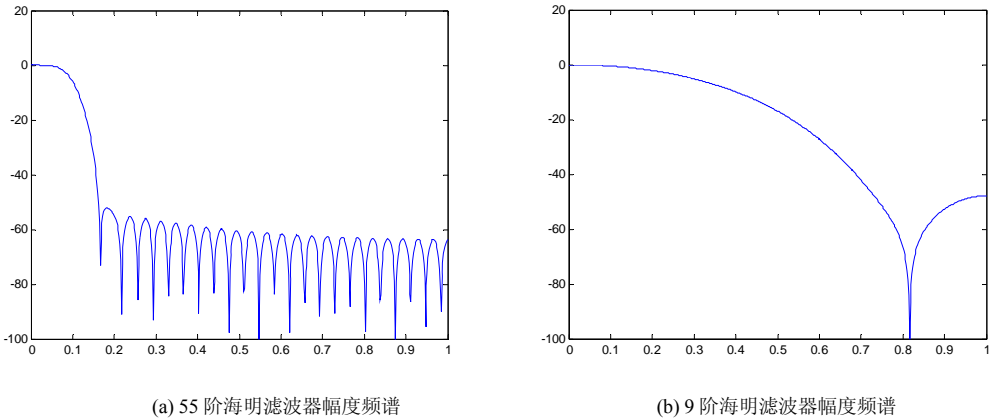


图 1 两组海明滤波器幅度频谱图

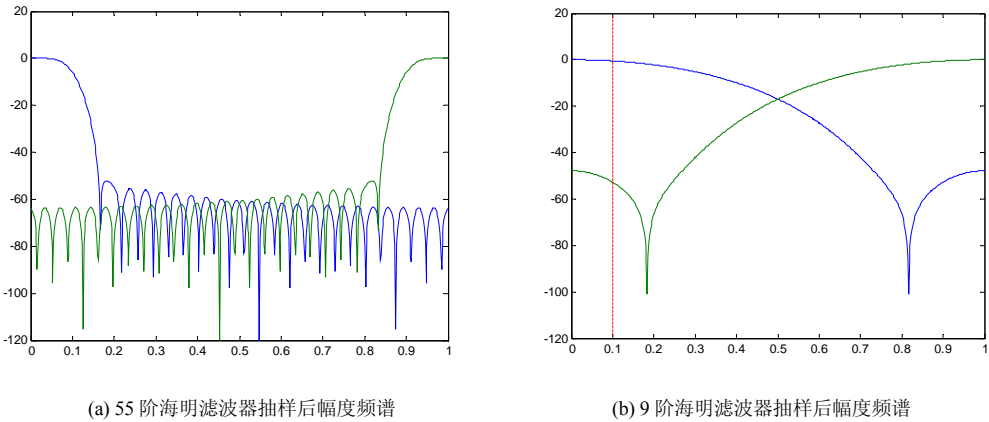


图 2 两组海明滤波器抽取后幅度频谱

通常，理想的方位预处理滤波器的频域特性为

$$H_d(e^{j\omega})=\begin{cases} e^{-j\omega\alpha}, & -\omega_c \leq \omega \leq \omega_c \\ 0, & \omega_c \leq \omega \leq 1, -1 \leq \omega \leq -\omega_c \end{cases} \tag{2}$$

而本文提出的二级滤波处理方法中，对第一级滤波器的理想频域特性要求可以表示为

$$H_d(e^{j\omega})=\begin{cases} e^{-j\omega\alpha}, & -\omega_c \leq \omega \leq \omega_c \\ 0, & 1-\omega_c \leq \omega \leq 1, -1 \leq \omega \leq -1+\omega_c \end{cases} \tag{3}$$

可以看出在 $(\omega_c,1-\omega_c)$ 和 $(-1+\omega_c,-\omega_c)$ 的区间里并没有提出要求，这相当于大大放松了对滤波器的设计要求。也为滤波器阶数降低的可能提供了理论依据。

从上面的分析可以看出，进行两级滤波需要一定的条件，一方面要求采样频率相对于信号带宽要足够大，能够满足低阶滤波处理和降采样要求，同时不对有效信号形成比较大的影

响；另一方面。要求在第二级滤波和后续处理中能够有效剔除有效信号频谱之外的无用信号。

3 两级滤波方法的实际应用

为了验证本文方法的有效性，我们结合实际 SAR 系统进行分析，某机载 SAR 系统，雷达的信号采样点数为 32K 字节，脉冲重复频率（PRF）为 2000Hz，考虑到雷达中频解调后为 I、Q 两路信号，SAR 雷达的持续数据率约为 128M 字节/秒。对这样高的数据流进行处理，对实时成像处理器带来比较高的要求。通常，SAR 实时成像处理器采用多级流水并行处理方案，大数据量、高数据率除了对数据运算、缓存提出要求外，对数据的传输也带来了不可忽视的瓶颈问题。为了降低数据量大带来的问题，可以采用两级滤波方法。首先用低阶数的滤波器进行滤波后降采样，减少一半的数据量，然后再进行后续的处理。由于数据率由 128M 字节/秒降低为 64M 字节/秒，对无论是运算、缓存还是传输的压力都大大降低，使 SAR 成像处理更容易实时实现。作为示例，我们选择 9 阶的布拉克曼滤波器进行第一级滤波。图 3(a)为布拉克曼滤波器的幅度频谱，图 3(b)为对应的滤波器 1/2 重抽样后的频域幅度谱。

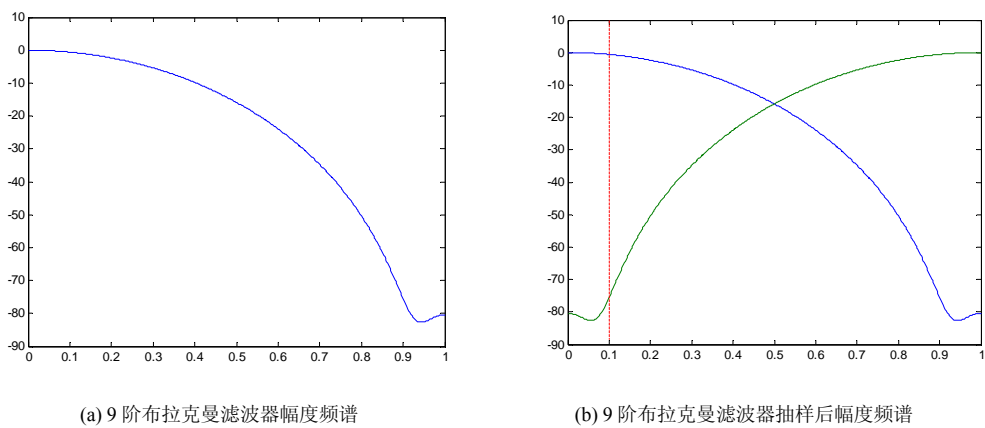


图 3 9 阶布拉克曼滤波频谱及抽取后幅度频谱

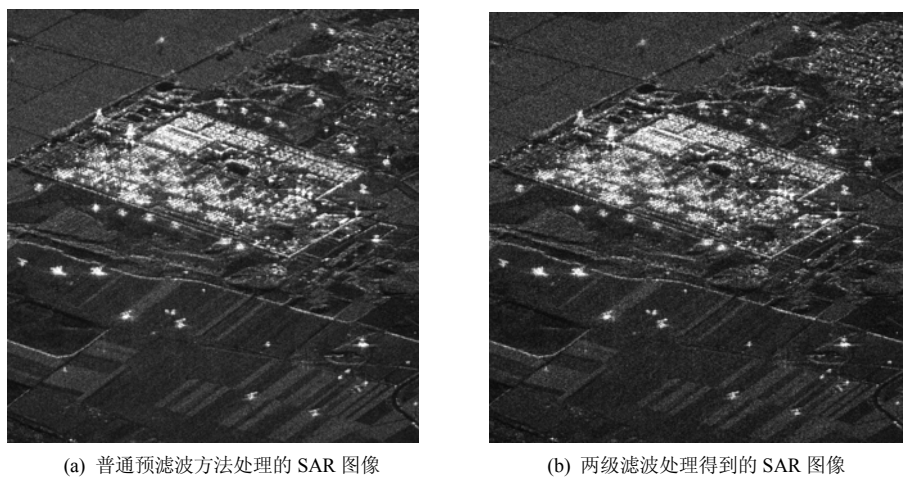


图 4 两种滤波方法处理得到的 SAR 图像

对于航空 SAR 系统而言,一般的脉冲重复频率远大于 SAR 雷达的多普勒频率带宽,如图 3(b)所示,经过滤波重抽样以后,在有效信号带宽以内的混叠影响甚至要小于图 2a 中的高阶数滤波器的混叠影响。后续的处理可以继续进行方位向的滤波处理,或者直接结合方位向的成像进行处理。作为示例,我们利用同一区域的 SAR 实际数据进行不同处理的对比分析。图 4(a)为采用普通预滤波方法得到的 SAR 图像,图 4(b)为采用两级滤波得到的 SAR 图像。可以看出,利用两种方法处理得到的图像与前面分析的一样,没有明显的差别。对这一实际数据的处理结果,也说明了本文所提出的两级滤波处理方法的正确性和有效性。

4 结束语

为了实现高性能的合成孔径雷达实时成像处理器,通常采用方位向预滤波降采样技术,在保证信号性能的同时减少数据量以降低对成像处理器的运算和存储等能力的要求。对 SAR 方位预处理的研究,通常是在限定滤波器阶数的前提下设计出性能良好的滤波器。本文针对实际系统的实现要求,提出了一种两级滤波降采样的实现方法。首先对雷达方位向信号进行低阶数的滤波处理,虽然滤波器过渡带比较宽,但可以保证有效信号的性能。在第一级滤波降采样后,可以减少至少一倍的数据量,降低对处理器的处理运算、缓存等要求的压力。之后进行第二级高性能滤波,可以提取出有效的信号,保证 SAR 雷达的成像质量。本文提出的方法适用于实时 SAR 成像处理器,特别是数据量和数据率比较大的场合,实际雷达数据的处理结果验证了本文方法的有效性。并且本文两级滤波的方法,也可以推广到多级滤波处理。本文中,对两级滤波器的设计提出了一般性的要求,在实际实现中需要结合具体的问题进一步细化设计。

参 考 文 献

- [1] Kirk J C, A Discussion of Digital Processing In Synthetic Aperture Radar, IEEE Transactions on Aerospace and Electronic Systems, Vol. AES-11(3),1975: 326-337.
- [2] Brown W M, Synthetic Aperture Processing With Limited Storage and Presumming, IEEE Transactions on Aerospace and Electronic Systems, Vol. AES-9(2),1975: 166-176.
- [3] Gudim M A, Cheng F, Madsen S, and etc, Single-chip High-Density FPGA Implementation of the SAR Azimuth Pre-filter for On-board Data Reduction, Earth Science Technology Conference 2002, Pasadena Hilton, Pasadena, CA, June.2002.
- [4] Curlander J C, McDonough R N, Synthetic Aperture Radar-Systems and Signal Processing, New York: John Wiley & Sons, 1991.
- [5] 程佩青. 数字信号处理教程. 北京: 清华大学出版社, 1995.

一种提高GNSS卫星信号模拟器伪距精度的方法

彭 鸣¹ 吴晓非² 邹永忠²

(1.北京邮电大学 信息与通信工程学院, 北京 100876; 2.北京邮电大学, 北京)

摘 要: 伪距精度是 GNSS 卫星信号模拟器的重要技术指标。而系统时钟的稳定度直接影响模拟器所能提供仿真信号的精度。本文提出一种在 GNSS 模拟器时钟采用普通晶振的条件下, 通过过采样内插成型滤波器的方法来提高模拟器的伪距精度。以一个 GPS 系统实例仿真与测试, 结果表明本文提出的方法使伪距精度提高了 10 倍。

关键词: GNSS 模拟器; 卫星信号; 内插成型滤波器; 伪距

A Method Of Improving GNSS Satellite Signal Simulator's Pseudo-range Precision

Peng Ming¹ Wu Xiao fei² Zou Yong zhong³

(1. School of Information and Telecommunication Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, China;

2. Beijing University of Posts and Telecommunications, Beijing 100876, China)

Abstract: The precision of the pseudo-range is an important parameter of the GNSS satellite signal simulator. The stability of the crystal oscillator influence simulator that can provide the simulation signal's precision and impact on the positioning result precision of the pseudo-range. Based on the common crystal oscillator of the GNSS simulator clock, this paper proposes a method that used over-sampling interpolation shaping filter to improving the simulator precision of the pseudo-range. According to this method by a GPS system simulating and testing, the result indicated that the precision of the pseudo-range improve 10 times.

Keywords: GNSS Simulator; Satellite signal; Interpolation shaping filter; Pseudo-range.

1 引言

全球导航卫星系统 (Global Navigation Satellite System--GNSS) 具有连续的全球覆盖、全天候、高精度、三维卫星导航定位的特性^[1]。该系统应用范围广泛, 如深空探测、陆海空导航、地质勘探、车辆定位系统等。目前, 一些国家都相继投资建设自己的卫星导航系统, 如美国的 GPS、俄罗斯的 GLONASS、欧盟的 GALILEO 和我国的北斗系统。GNSS 卫星信号模拟器就是模拟产生各种实际环境下的 GNSS 卫星信号^[2]。根据载体动态特性等各种因素的影响, 精确产生 GNSS 接收机接收到的各颗 GNSS 卫星信号, 充分模拟载体运动过程中 GNSS 接收机遇到的各种复杂场景, 为 GNSS 接收机的研制开发、测试提供真实的仿真环境。

目前, Spirent、安捷伦公司研制的模拟器能模拟多种运动场景, 满足多频点、多通道、高精度的需求, 但对于一些开发 GNSS 接收机的中小企业来说, 购买 Spirent、安捷伦公司生产的高精度时钟信号模拟器成本较高, 而自主研发的模拟器, 由于系统时钟采用的是普通晶振(频偏在 1ppm-10ppm)会对伪距精度造成 $10^2 \sim 10^5$ m 的误差。另一方面, 对于卫星发射出来的 GNSS 信号, 接收机能够通过安装天线接收 GNSS 信号, 但它不能根据测试需要来控制模拟器发出各种运动场景的 GNSS 信号, 具有不可重复性。在这样的条件下, 本文提出一种过采样成型滤波器来提高模拟器伪距精度的方法, 文中以美国 GPS(Global Positioning System—全球定位系统)卫星信号模拟器为例, 对过采样成型滤波器在 GPS 系统中进行设计, 性能仿真与评估结果, 得出结论。

2 过采样成型滤波器的设计

由于卫星和接收机的时钟一般不是同步的, 它们分别与系统时之间都有偏移误差。因此, 实际的观测距离并不是真实的卫星到用户的距离, 而存在一个偏移, 将这一观测量称为伪距 ρ , 它是通过将信号传播速度乘以两个非同步时钟(卫星钟和接收机钟)之间的时间差而确定的距离。伪距包含从卫星到用户的真实距离; 由系统时与用户时钟之间的差异而造成的偏移; 系统时和卫星时钟之间的偏移。伪距的计算公式如下:

$$\rho = c[(T_u + t_u) - (T_s + \delta)] \quad (1)$$

式中: T_s 表示信号离开卫星时的系统时; T_u 表示信号到达用户接收机时的系统时; δ 表示卫星时钟与系统时之间的偏移, 超前为正, 滞后(延迟)为负; t_u 表示接收机时钟与系统时之间的偏移; $T_s + \delta$ 表示在信号离开卫星时卫星时钟的读数; $T_u + t_u$ 表示信号到达用户接收机时的用户接收机时钟的读数; c 表示光速。

伪距精度与卫星时钟(模拟器时钟)和接收机时钟有关, 本文考虑模拟器采用普通晶振条件下的时钟偏移。由原子钟驱动的系统时, 偏差是在 50ns 以内, 采用普通晶振的模拟器频偏在 1~10ppm。在时钟采用普通晶振的情况下, 提出一种过采样成型滤波器来提高模拟器伪距精度的方法。

过采样成型滤波器采用 N 阶、L 倍内插 FIR 滤波器。在模拟器的发送端, 将数据流信号经过成型滤波器内插 L 倍, 在经过上变频将信号发送出去。经过成型滤波器的数据流速率提高了 L 倍, 以补偿石英晶振时钟带来的偏差, 也增加 D/A 转换后的精度。L 倍内插的成型滤波器采用查找表的算法, 还可以为 FPGA 硬件系统中节省大量的乘法器和多级累加器。而没有内插的成型滤波器, 数据流只是经过一个脉冲成型滤除带外干扰, 数据流的速率没有提高, 不会对结果带来明显的提高。

式(1)中伪距与模拟器的时钟可以看成线性变化, 理论分析内插 L 倍的数据流, 通过模拟器产生的信号在基带板接收处理得到的伪距精度会提高 L 倍, 后面会列举一个实例进行分析。对于成型滤波器的阶数 N 和内插系数 L 可以根据具体系统的参数指标来确定。

我们采用升余弦滚降滤波器作为成型滤波器, 扩频后的数据经过成型滤波器, 也就是对基带信号进行脉冲成型滤波, 限制发送信号的带宽, 滤除掉带外干扰。滤波器的输出表示为:

$$y(n) = h(n) * x(n) = \sum_{k=0}^{N-1} h(k)x(n-k) \quad (2)$$

对基带数字信号进行 L 倍零内插，即每相邻两输入序列点之间内插 $L-1$ 个零点。内插后信号频谱为原始序列谱经 L 倍压缩后得到的频谱。相邻两输入序列间内插 $L-1$ 个零的输入为：

$$v(n)=\begin{cases} x(n/L) & n=0,\pm L,\pm 2L,\cdots \\ 0 & \text{其他} \end{cases} \tag{3}$$

输出表示为：

$$y(n)=v(n)*h(n)=\sum_{k=0}^{\infty} v(n-k)h(k)=\sum_{k=0}^{\infty} x(n-Lk)h(k) \quad (n=0,\pm L,\pm 2L,\cdots) \tag{4}$$

由上式看出输出的 $y(n)$ 是通过冲激响应 $h(n)$ 与数据流 $x(n)$ 相乘累加得到，这样做在 FPGA 硬件系统中需要用到大量的乘法器和多级累加器，会占用大量的资源。为了节省资源，我们还可以采用基于 ROM 查找表的算法，将相乘累加后的值存储在 ROM 表中，根据数据流对应的地址来查找相应的输出值，这样可以减少一定的电路资源，并提高系统的处理速度。

3 具体应用实例设计实现

由于美国的 GPS 系统已经是全运行的，本文以 GPS 系统为例来说明过采样成型滤波器在提高伪距精度中的具体应用。GPS 卫星信号模拟器的工作原理见图 1 所示。

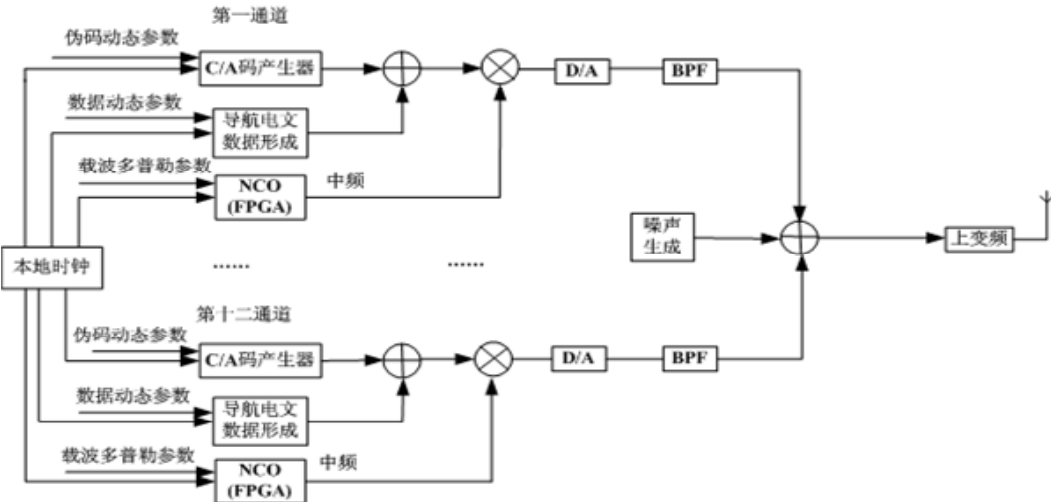


图 1 GPS 卫星信号模拟器组成原理

图中 BPF 为带通滤波器，通过带通滤波器滤出所需的中频信号。GPS 卫星信号模拟器就是要产生 $L1$ 频率的 C/A 码信号，在用户天线端可以接收到这个可见的 GPS 卫星信号。由于 $P(Y)$ 码信号保密，因此我们研制的信号模拟器仅模拟 $L1$ 频率(1575.42MHz)的 C/A 码信号，这些信号的表达式为：

$$S(t)=\sum_i A_i(t)C_i(t-\tau_i(t))D_i(t-\tau_i(t))\sin(2\pi f_i(t-\tau_i(t))+\varphi_i(t)) \tag{5}$$

式中： $A_i(t)$ 表示 t 时刻第 i 颗卫星的信号幅度； $C_i(t)$ 表示 t 时刻第 i 颗卫星发射的 C/A 码信号； $D_i(t)$ 表示在第 i 颗卫星信号上调制的导航电文数据； $f_i(t)$ 表示第 i 颗卫星的瞬时载波频率； $\varphi_i(t)$ 表示载波信号的相位； τ_i 表示第 i 颗卫星信号传播的群延迟。

GPS 能定位成功的关键就在于有高稳定度的频率标准和高稳定度的时间系统。时钟是整个模拟器的重要组成部分，是影响模拟器精度的关键因素。模拟器必须保证提供高精度的时钟信号，为整个软件系统提供精确的 GPS 时。但是一般模拟器大都采用的是石英晶振，很少会用到像原子钟的高稳晶振，就需要在普通晶振的条件下，从其它方面改善，减少对系统造成误差。

本文提出在时钟采用普通晶振的条件下，通过加入过采样成型滤波器来提高伪距精度的方法。根据硬件系统的参数指标，过采样成型滤波器确定采用 64 阶、16 倍内插的升余弦滚降系数数字滤波器就能够达到提高 10 倍的要求。

在图 1 中，导航电文和 C/A 码信号调制形成数据流信息，将此数据流通过成型滤波器，D/A 转换，再经过上变频为中频信号发送出去。可以对经过成型滤波器的数据流进行 ModelSim 仿真，数据流经过成型滤波器后位数增加，提高了 D/A 转换后的精度。在用户接收端接收模拟器发出的中频模拟信号，进行下变频、捕获、跟踪得到伪距值，并对伪距进行解算得到用户的位置信息。可以在硬件基带板上进行测试，观察加入成型滤波器后，伪距精度的变化。

采用 10.23MHz 时钟信号作为基准时钟频率，完成扩频调制后进行 16 倍零内插，结果送给 D/A 转换器，经过数字上变频来实现中频信号的频率变换，得到 GNSS 模拟器中频模拟信号。实现原理图见图 2 所示：

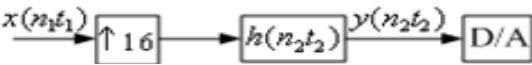


图 2 GPS 模拟器基带/中频信号变换原理

4 性能仿真与评估结果

通过对 GPS 系统中加入过采样 64 阶、16 倍内插、升余弦滚降系数的成型滤波器模型的讨论，并进行 ModelSim 仿真和基带板测试。图 3 是数据流经过成型滤波器后的 ModelSim 仿真波形图。经过成型滤波器的数据流并不是直接与滤波器的冲激响应相乘累加，而是将相乘累加后的结果存储在 ROM 表中，通过查找表的方式输出。

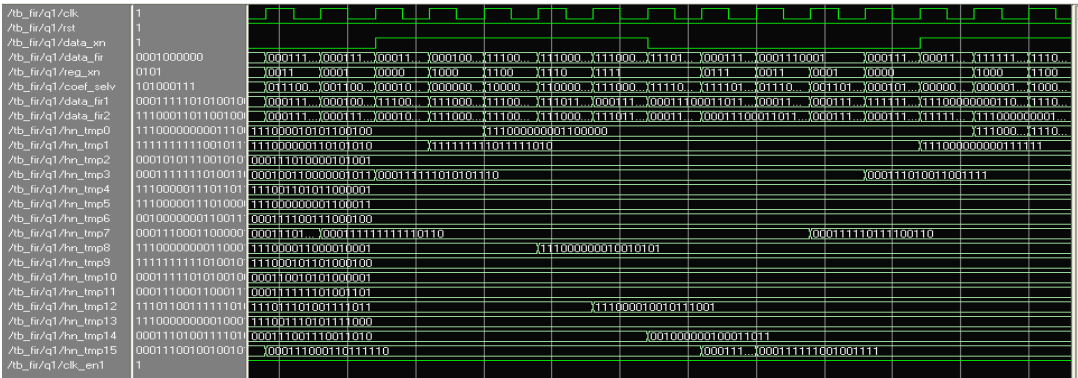


图 3 扩频后过采样成型滤波器的仿真波形

以上只是加入成型滤波器后，对程序设计部分的一个软件仿真实现，还需要通过硬件部分的基带板测试，更加直观的反应伪距精度的提高。将代码综合仿真生成能基带板上能够测

试的可执行文件，在硬件基带板上进行测试，得到伪距值进行坐标值变换，得到用户的三维位置，也就是对应的经度、纬度和高度值。选取其中的 10 组数据，比较加入成型滤波器后的变化如表 1 所示。

表 1 加入过采样成型滤波器前后的比较

加过采样成型滤波器方案前			加过采样成型滤波器方案后		
经度	纬度	高度	经度	纬度	高度
39.999684	115.999284	+94.114035	40.0001111	115.9999513	-00.343212
39.999702	115.999206	+90.667645	40.0000835	115.9999511	+03.965101
39.999721	115.999260	+84.776281	40.0001135	115.9999816	-03.870483
39.999699	115.999322	+88.320528	40.0001073	116.0000377	-05.329556
39.999720	115.999221	+84.670087	40.0001121	116.0000049	-07.196734
39.999672	115.999232	+95.993096	40.0000538	115.9999662	+03.629919
39.999689	115.999275	+89.802758	40.0000756	115.9999334	+02.703299
39.999731	115.999266	+78.643199	40.0000803	115.9999989	+00.304294
39.999666	115.999269	+95.539130	40.0001221	116.0000256	-09.020110
39.999682	115.999212	+92.677109	40.0001025	116.0000289	-03.272981

模拟器所采用的本地时钟频偏在 1~10ppm，在不考虑跟踪环路误差和采样偏差的条件下，时钟对伪距造成几十米的偏差。通过测试结果加入成型滤波器后，高度由 90m 降到-10~10m 的波动，如图 4 所示。可见加入成型滤波器后高度有明显变化，取 100 组高度的数据结果进行比较。

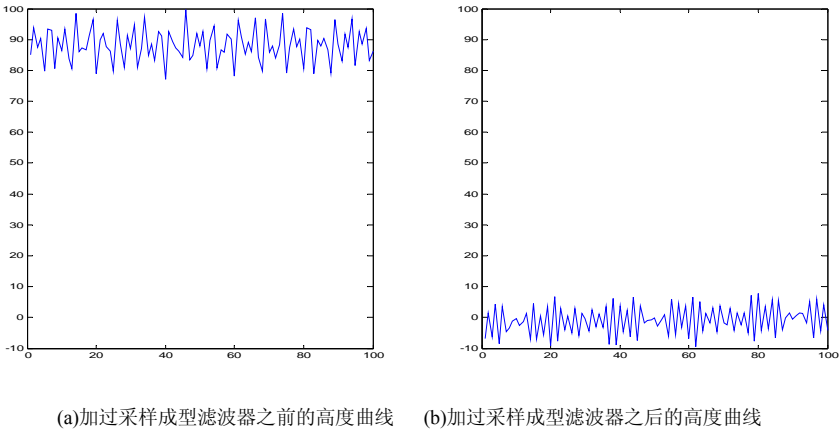


图 4

由以上数据和波形图可见，加入过采样成型滤波器之后，提高了伪距的精度，测试结果在-10~10m 内波动。

5 结论

通过 ModelSim 仿真和基带板测试，采用普通晶振（精度在 1~10ppm）作为本地时钟，

不考虑跟踪环路误差和采样偏差的情况下,测试结果伪距高度有 90m 的偏差,加入过采样成型滤波器后,伪距的高度在-10~10m 波动。本文讨论加入过采样 N 阶、L 倍内插的升余弦滚降系数成型滤波器来提高伪距精度的方法,并用一个具体实例设计实现,对于滤波器的阶数和内插系数可以根据具体系统来确定。在这个设计实例中过采样成型滤波器采用的是 64 阶、内插系数为 16 倍的升余弦滚降系数数字滤波器,可以看出测试的结果提高了 10 倍。在此基础上还可以提高成型滤波器的阶数和内插系数,使伪距的高度进一步提高。此方法对于俄罗斯的 GLONASS、欧盟的 GALILEO 系统和北斗系统,稍加改动也同样适用。

参 考 文 献

- [1] Elliott D.Kaplan, Christonpher J.Hegarty 等主编;寇艳红译. GPS 原理与应用(第二版)[M]. 北京:电子工业出版社,2007.7.
- [2] 赵军祥,张其善,常青,等. 高动态 GPS 卫星信号模拟器关键技术分析及应用[J]. 电视技术,2003(4): 49-54.
- [3] 李宝柱,张其善,杨东凯. GPS 卫星信号模拟器中频信号处理与实现[J]. 北京航空航天大学学报,2006, 32(07): 811-814.
- [4] 贾科军,柯熙政. 高动态 GPS 卫星信号模拟器设计与实现[J]. 宇航计测技术,2006, 26(5): 1-6.
- [5] 胡广书编著. 数字信号处理-理论、算法与实现(第二版)[M]. 北京:清华大学出版社,2003.8.
- [6] 张伯川,常青,张其善,等. 高动态 GPS 卫星信号模拟器导航电文生成[J]. 北京航空航天大学学报,2005, 31(3): 284-287.
- [7] 胡修林,杨志专,张蕴玉. 数字中频 GPS 信号软件模拟器设计[J]. 电视技术,2007, 47(01): 108-111.
- [8] 张颖慧,常青,张其善,等. 高动态 GPS 信号模拟器的软件实现[J]. 无线电工程,2003, 33(07): 14-17.

相干信源的二维DOA盲估计算法

余 俊 陈未央 张小飞

(南京航空航天大学 电子工程系, 江苏省 南京市 210016)

摘 要: 相干信源 DOA 估计是阵列信号处理的一个研究热点。空间平滑技术是目前一种较有效的解相干处理算法,但它是通过牺牲阵列有效阵元数来换取解相干能力的,阵列孔径有一定的损失。本文通过构建二个 Toeplitz 矩阵,使接收相干信源协方差矩阵恢复为满秩矩阵,从而达到去相干的目的,然后再利用传统的 ESPRIT 算法通过两步一维搜索进行二维波达方向估计,仿真结果表明此算法有较好的二维相干信号的 DOA 估计性能,和算法的分辨性能也优于常规的空间平滑技术,而且该算法无需谱峰搜索,具有较低的计算量。

关键词: 相干信号; DOA 估计; ESPRIT 算法; 十字阵列; Toeplitz 矩阵

Blind 2-D DOA Estimation algorithm for Coherent Signal

Yu Jun Chen Wei yang Zhang Xiao fei

(Electronic Engineering Department, Nanjing University of aeronautics & astronautics, Nanjing, China 210016)

Abstract: The method of coherent DOA estimation is a research hotspot of array signal processing. Currently, the spatial smoothing algorithm is a more effective way, whose ability of decorrelation is gained by expense of the number of arrays element, but there is a certain degree of loss of the array aperture. In this paper, a 2-D coherent DOA estimation method using ESPRIT algorithm by Two-step one-dimensional search is proposed, which reconstructs a vertical crossed array model based on two Toeplitz matrices whose rank is only related to the DOA of signals and cannot be affected by the coherency between them, Simulation results show a better performance that this algorithm is far superior to conventional spatial smoothing techniques, which has lower computational complexity and without using spectral peak searching algorithm.

Keywords: Coherent signals; direction-of-arrival (DOA) estimation; ESPRIT algorithm; vertical crossed array; Toeplitz matrices.

1 引言

空间信号的波达方向(DOA, Direction Of Arrival)估计问题在阵列信号处理研究中具有重要意义,常用的 DOA 估计方法有谱分析、最大似然法和子空间方法等^[1-6]。但当信号源相干时,阵列接收的数据协方差矩阵不是满秩,这就会导致信号子空间的维数小于信号源数,也就是说信号子空间“扩散”到了噪声子空间,从而子空间方法不能准确估计来波方向。解决思

路围绕着如何补偿由于相干引起的秩亏缺。为此，人们提出了很多算法，传统解相干主要是一些空间平滑算法^[7-13]，但这个过程通常只适用于等距均匀线阵，而且修正后矩阵的维数小于原矩阵的维数，也就是说解相干性能是通过降低自由度换取的。空间平滑算法的实质是通过降低矩阵的维数来恢复数据协方差矩阵的秩，从而达到解相干的目的。

文献[14]在解决矩阵秩亏损的问题上提出了一种新的 ESPRIT 方法，即通过特殊的天线阵列模型，重构一个 Toeplitz 矩阵，使其秩只与信号的波达方向有关，而不受信号相关性的影响，从而达到去相关的目的。文献[15]在[14]的基础上提出三线性分解的方法，并使得性能大为提高。文献[14, 15]只能估计一维角度，本文将在此基础上进行研究，利用互相垂直的两个均匀线阵及信号子空间与方向向量之间的关系，运用经典的 ESPRIT 算法，通过两步一维搜索估计二维波达方向，实现了相干信源的二维到达角方向估计，且该算法无需谱峰搜索，具有较低的计算量。

本文第 2 节给出了数据模型；第 3 节提出 2D-DOA 估计的 ESPRIT 算法；第 4 节通过计算机仿真证明了算法的性能改善情况；第 5 节对该算法进行了概括总结。

2 数据模型

阵列是一个十字阵（如图 1 所示），设十字阵的交点为坐标原点，所有阵元均匀分布在 x 轴和 y 轴上， x 轴和 y 轴上的阵元数均为 $N=2M+1$ ，共有 $4M+1$ 个阵元。阵元间距为 d 。假定有 P 个信号入射到该系统上，其中前 L 个为相干信号，其他为非相干信号并且独立与前 L 个信号。第 i 个信号的二维到达角分别为 θ_i, ψ_i ($i=1,2,\dots,P$)， θ_i, ψ_i 分别为方位角和俯仰角，

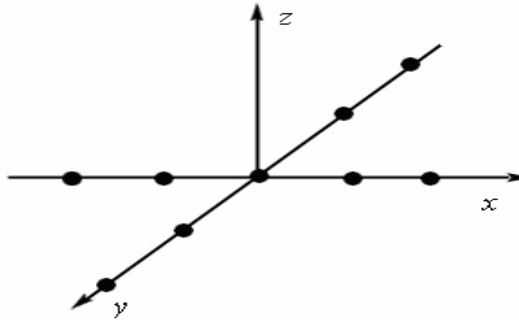


图 1 阵列的结构

则 x 轴和 y 轴的阵列接收信号为：

$$\mathbf{X} = \mathbf{A}_x \mathbf{S} + \mathbf{N}_x \quad (1.1)$$

$$\mathbf{Y} = \mathbf{A}_y \mathbf{S} + \mathbf{N}_y \quad (1.2)$$

其中， \mathbf{X} 和 \mathbf{Y} 分别为 x 轴和 y 轴接收快拍数据矩阵， \mathbf{A}_x 和 \mathbf{A}_y 为阵列的方向矩阵， \mathbf{S} 为信源矩阵， $\mathbf{S} = [\mathbf{s}_1 \ \cdots \ \mathbf{s}_1 \ \mathbf{s}_{L+1} \ \cdots \ \mathbf{s}_P]$ ，接收阵列协方差 $\mathbf{R}_{xx} = \mathbf{X}\mathbf{X}^H$ ， $\mathbf{R}_{yy} = \mathbf{Y}\mathbf{Y}^H$ ，根据参考文献[14]，协方差函数 \mathbf{R}_{xx} ， \mathbf{R}_{yy} 可以表示如下：

$$r_x(m, n) = \sum_{i=1}^P d(x)_{m,i} e^{j2\pi d n \sin \theta_i \sin \psi_i / \lambda} + \sigma_{nx}^2 \delta_{m,n}, m, n = -M, \dots, 0, \dots, M \quad (2.1)$$

$$r_y(m, n) = \sum_{i=1}^P d(y)_{m,i} e^{j2\pi dn \cos \theta_i \sin \psi_i / \lambda} + \sigma_{ny}^2 \delta_{m,n}, m, n = -M, \dots, 0, \dots, M \quad (2.2)$$

其中 $d(x)_{m,i}, d(y)_{m,i}$ 见文献[14], 重构二个 Toeplitz 矩阵得:

$$\mathbf{R}_x(m) = \begin{bmatrix} r_x(m, 0) & r_x(m, 1) & \cdots & r_x(m, M) \\ r_x(m, -1) & r_x(m, 0) & \cdots & r_x(m, M-1) \\ \vdots & \vdots & \ddots & \vdots \\ r_x(m, -M) & r_x(m, -M+1) & \cdots & r_x(m, 0) \end{bmatrix} \in \mathbb{C}^{(M+1) \times (M+1)}, m = -M, \dots, 0, \dots, M \quad (3.1)$$

$$\mathbf{R}_y(m) = \begin{bmatrix} r_y(m, 0) & r_y(m, 1) & \cdots & r_y(m, M) \\ r_y(m, -1) & r_y(m, 0) & \cdots & r_y(m, M-1) \\ \vdots & \vdots & \ddots & \vdots \\ r_y(m, -M) & r_y(m, -M+1) & \cdots & r_y(m, 0) \end{bmatrix} \in \mathbb{C}^{(M+1) \times (M+1)}, m = -M, \dots, 0, \dots, M \quad (3.2)$$

$$= \mathbf{A}_{rx} \mathbf{D}(m)_x \mathbf{A}_{rx}^H + \mathbf{N}_{mx}$$

其中,

$$\mathbf{A}_{rx} = [\mathbf{a}_{rx}(\theta_1) \quad \mathbf{a}_{rx}(\theta_2) \quad \cdots \quad \mathbf{a}_{rx}(\theta_P)] \quad (4.1)$$

$$\mathbf{A}_{ry} = [\mathbf{a}_{ry}(\theta_1) \quad \mathbf{a}_{ry}(\theta_2) \quad \cdots \quad \mathbf{a}_{ry}(\theta_P)] \quad (4.2)$$

$$\mathbf{a}_{rx}(\theta_p) = \begin{bmatrix} 1 & e^{-j\frac{2\pi}{\lambda}d \sin \theta_p \sin \psi_p} & \cdots & e^{-j\frac{2\pi}{\lambda}dM \sin \theta_p \sin \psi_p} \end{bmatrix}^T \quad (5.1)$$

$$\mathbf{a}_{ry}(\theta_p) = \begin{bmatrix} 1 & e^{-j\frac{2\pi}{\lambda}d \cos \theta_p \sin \psi_p} & \cdots & e^{-j\frac{2\pi}{\lambda}dM \cos \theta_p \sin \psi_p} \end{bmatrix}^T \quad (5.2)$$

$$\mathbf{D}(m)_x = \text{diag}\{d(x)_{m,1} \quad d(x)_{m,2} \quad \cdots \quad d(x)_{m,P}\} \quad (6.1)$$

$$\mathbf{D}(m)_y = \text{diag}\{d(y)_{m,1} \quad d(y)_{m,2} \quad \cdots \quad d(y)_{m,P}\} \quad (6.2)$$

因为 $\mathbf{A}_{rx}, \mathbf{A}_{ry}$ 是一个范德蒙矩阵, 且 $\theta_i \neq \theta_j$, 所以 $\mathbf{A}_{rx}, \mathbf{A}_{ry}$ 满秩, 又因为 $\mathbf{D}(m)_x, \mathbf{D}(m)_y$ 是一个对角矩阵, 对 $\mathbf{D}(m)_x, \mathbf{D}(m)_y$ 进行特征值分解, 我们得到 P 个大特征值和 $M-P+1$ 个小特征值, 其中大特征值对应的大特征向量构成信号子空间 $\Phi_s = \text{span}\{v_1 \quad v_2 \quad \cdots \quad v_P\}$, 其中小特征值对应的小特征向量构成噪声子空间 $\Phi_n = \text{span}\{v_{P+1} \quad v_{P+2} \quad \cdots \quad v_{M+1}\}$ 。

3 相干信源 2D-DOA估计

$\mathbf{A}_{rx} = \begin{bmatrix} a_1 \\ \mathbf{A}_{Bx} \end{bmatrix} = \begin{bmatrix} \mathbf{A}_{Fx} \\ a_{M+1} \end{bmatrix}$, 其中 a_1, a_{M+1} 为 \mathbf{A}_{rx} 的第一列和最后一列。同样, 我们对信号子空间进行分解, $\{v_1 \quad v_2 \quad \cdots \quad v_P\} = \mathbf{U}_{Sx} = \begin{bmatrix} u_1 \\ \mathbf{U}_{Bx} \end{bmatrix} = \begin{bmatrix} \mathbf{U}_{Fx} \\ u_{M+1} \end{bmatrix}$, 其中, u_1, u_{M+1} 为 \mathbf{U}_{Sx} 的第一列和最后一列。

易知: $\mathbf{A}_{Bx} = \mathbf{A}_{Fx} \Phi_x$, 其中,

$$\Phi_x = \text{diag} \left\{ e^{-j\frac{2\pi}{\lambda}d \sin \theta_1 \sin \psi_1}, e^{-j\frac{2\pi}{\lambda}d \sin \theta_2 \sin \psi_2}, \dots, e^{-j\frac{2\pi}{\lambda}d \sin \theta_P \sin \psi_P} \right\} \quad (7)$$

称为旋转矩阵，因为 A_{rx} 和 U_{sx} 张成同样的信号子空间，所以：

$$U_{Fx} = A_{Fx}T, U_{Bx} = A_{Bx}T \quad (8)$$

其中 T 为满秩矩阵，

$$U_{Fx}T^{-1}\Phi_xT = A_{Fx}TT^{-1}\Phi_xT = A_{Fx}\Phi_xT = A_{Bx}T = U_{Bx} \quad (9)$$

令 $\psi_x = T^{-1}\Phi_xT$ ，有

$$U_{Fx}\psi_x = U_{Bx}, \quad \psi_x = [U_{Fx}]^+ U_{Bx} \quad (10)$$

对 ψ_x 进行特征值分解我们可以得出 $e^{-j\frac{2\pi}{\lambda}d \sin \theta_k \sin \psi_k}$ ， $k=1,2,\dots,P$ ，即可估计出 $\sin \theta_k \sin \psi_k$ ，同理我们可以求得 ψ_y ，及估计 $\cos \theta_k \sin \psi_k$ ，通过配对，二维方位角估计即可利用下面公式求得：

$$\theta_k = \tan^{-1} \left(\frac{\sin \theta_k \sin \psi_k}{\cos \theta_k \sin \psi_k} \right) \quad (11.1)$$

$$\psi_k = \sin^{-1} \left(\sqrt{(\sin \theta_k \sin \psi_k)^2 + (\cos \theta_k \sin \psi_k)^2} \right) \quad (11.2)$$

为此我们提出了一种基于 ESPRIT 的相干信源的 2D-DOA 估计算法，其算法步骤如下：

1. 由两个子阵的接收数据 X, Y ，由式 (2.1)，(2.2) 得到数据协方差矩阵 $r_x(m, n), r_y(m, n)$ ；
2. 由式 (3.1)，(3.2) 构造二个 Toeplitz 矩阵；
3. 对构造的 Toeplitz 矩阵进行特征值分解，得到信号子空间 U_s ；
4. 对 U_s 进行分解，得到 U_F, U_B ，按式 (10) 得到矩阵 ψ_x, ψ_y ，然后对其进行特征值分解；
5. 通过配对，按照式 (11.1)，(11.2) 即可得到对应 P 个信号的二维到达角。

我们提出的算法该算法无需谱峰搜索，具有较低的计算量。它的复杂度主要集中在协方差矩阵的构造、特征数值分解和矩阵求逆。我们算法的复杂度为 $O(2(M+1)^3(2M+1) + 2(M+1)^3 + 2P^3)$ 。

4 算法仿真

仿真中采用三个信号源，其中前两个信号完全相干，三个信号的波达方向分别为：(10°, 15°)，(20°, 25°)，(30°, 35°)，X 轴和 Y 轴上的阵元数均为 13，间距为半波长，快拍数为 100。

图 2(a)和(b)分别给出了 SNR=25dB 和 SNR=40dB 下信号波达方向的估计分布情况，仿真结果说明：此算法有较好的二维相干信号的 DOA 估计性能，在信噪比达到 40dB 时能精确估计二维角度。

图 3(a)和(b)分别给出了 ESPRIT 算法和平滑算法方位角和俯仰角估计性能对比，我们定义： $MSE = \frac{1}{P \times 1000} \sum_{m=1}^{1000} \sum_{p=1}^P |\hat{\theta}_{m,p} - \theta_p|^2$ ，其中 $\hat{\theta}_{m,p}$ 是第 m 次仿真中第 p 信源的 DOA 估计值， θ_p 为第 p 信源的 DOA 精确值。从图 3 可以看出，本文算法的分辨性能远远优于常规的空间平滑技术。

图 4(a)和(b)分别给出了天线数分别为 11、13、15 时方位角和俯仰角估计性能对比，可以

看出，算法的性能随着天线数的增加，性能变好。

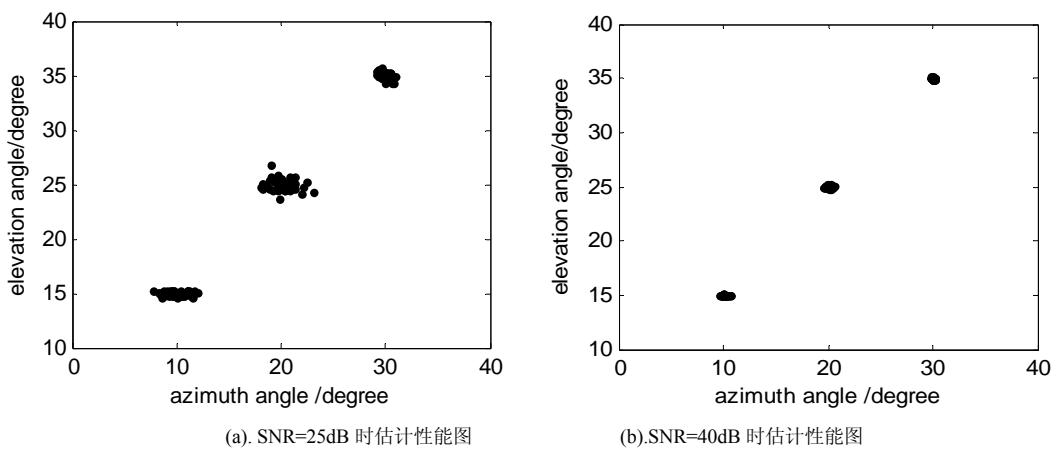


图 2

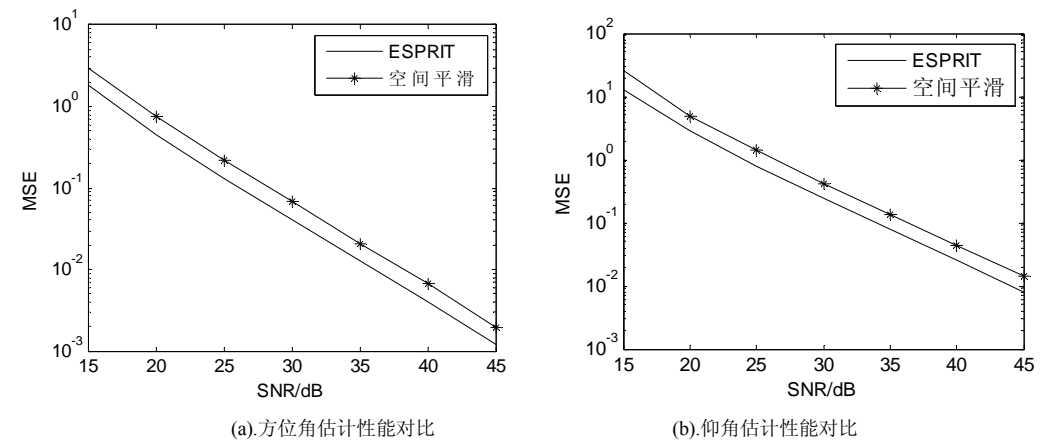


图 3

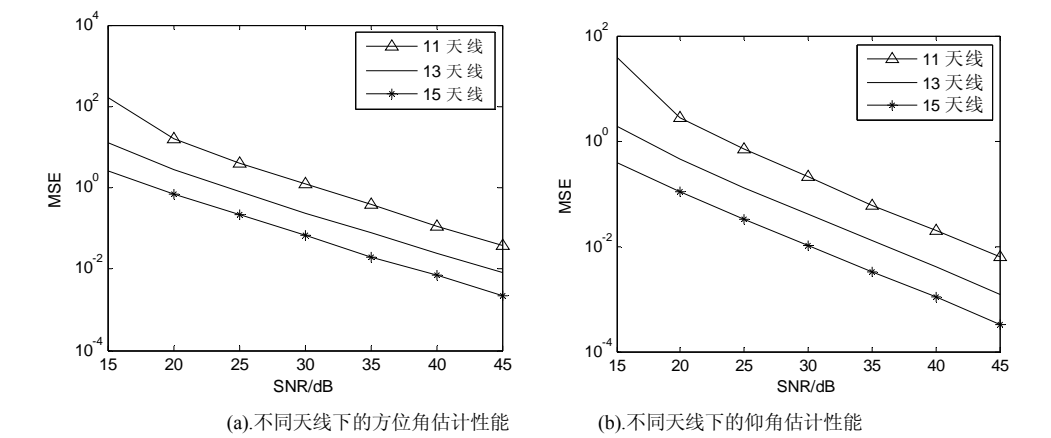


图 4

5 结论

本文主要研究了基于十字阵列的相干信源的 DOA 估计算法,该算法通过构建一个十字阵列和二个 Toeplitz 矩阵,用较简单的算法解决了以往解相干信源 DOA 估计的算法存在较大的阵列冗余度的问题,针对不同信噪比,不同阵列数进行了仿真分析,并且对不同的算法进行了仿真比较,仿真结果说明:此算法有较好的二维相干信号的 DOA 估计性能,算法的分辨性能优于常规的空间平滑技术,而且该算法无需谱峰搜索,具有较低的计算量。

参 考 文 献

- [1] Mendoza-Montoya F, Covarrubias-Rosales D.H, Lopez-Miranda C.A. DOA Estimation in Mobile Communications System using Subspace Tracking Methods [J]. *IEEE Latin America Transactions*, 2008, 6(2): 123 - 129.
- [2] Gotsis K.A., Siakavara K., Sahalos J.N. On the Direction of Arrival (DOA) Estimation for a Switched-Beam Antenna System Using Neural Networks [J]. *IEEE Trans. on Antennas and Propagation*, 2009, 57(5): 1399 - 1411.
- [3] Peng Zhenzhe, Wan Fangjie. A Novel DOA Algorithm for CDMA System [C]. *2009 International Conference on Networks Security, Wireless Communications and Trusted Computing*, 2009, 2: 32 - 35.
- [4] Rubsamen M., Gershman A.B. Direction-of-Arrival Estimation for Nonuniform Sensor Arrays: From Manifold Separation to Fourier Domain MUSIC Methods [J]. *IEEE Trans. on Signal Processing*, 2009, 57(2): 588 - 599.
- [5] Wu J.X., Wang T., Suo Z.Y.. DOA estimation for ULA by spectral Capon rooting method [J]. *Electronics Letters*. 2009, 45(1): 84 - 85.
- [6] Wen-Jun Zeng, Xi-Lin Li, Xian-Da Zhang. Direction-of-Arrival Estimation Based on the Joint Diagonalization Structure of Multiple Fourth-Order Cumulant Matrices [J]. *IEEE Signal Processing Letters*, 2009, 16(3): 164 - 167.
- [7] T J Shan, Vax M, on spatial smoothing for direction-of-arrival estimation of coherent signals[J], *IEEE Trans*, Aug, 1985, 33(4): 806-511.
- [8] 董玫,张守宏,吴向东,一种改进的空间平滑算法,电子与信息学报[J],2008,30(4): 859-862.
- [9] 齐崇英,王永良,张永顺,色噪声背景下相干信源DOA估计的空间差分平滑算法[J],电子学报, 2005, 33(7):1314-1318
- [10] Pillai S.U. Kwon B.H. Forward/backward spatial smoothing techlnques for coherent singals identification[J].*IEEE Trans*, on ASSP, 1989, 37(1): 8-15.
- [11] Kun DU D. Modified MUSIC algorithm for estimation DOA of Singals [J]. *Singal Poreessing*, 1996, (48): 85-891.
- [12] Wang H., Kaveh M. On the perfomance of singal-subspace processing-Part11: Cohernt wide-band systems [J], *IEEE Trans. On ASSP*, 1987, 35(11): 1583-1591
- [13] 王布宏,王永良,陈辉. 一种新的相干信源DOA估计算法: 加权空间平滑协方差矩阵的Toeplitz矩阵拟合[J], 电子学报 2003, 31 (9): 1394-1397
- [14] Fang-Ming Han, Xian-Da Zhang. An ESPRIT-like algorithm for coherent DOA estimation [J], *IEEE Antennas and Wireless Propagation Letters*, 2005, 4(1) : 443- 446.
- [15] Xiaofei Zhang , Dazhuan Xu. Improved Coherent DOA Estimation algorithm for Uniform Linear Array[J], *International Journal of Electronics*, 2009 , 96(2) : 213 - 222.

一种新的低功率信号检测算法

王磊 郑宝玉 崔景伍

(南京邮电大学 信号处理与传输研究院, 江苏, 南京 210003)

摘要: 以实现未来绿色通信为目标, 对如何运用自由概率理论(FPT)进行通信系统中的低功率信号检测进行了研究。FPT 是一种非常有用的用于研究系统渐近特性的数学工具。运用 FPT 方法可以从混有噪声的系统样本协方差矩阵入手, 用乘法自由解卷积表示样本协方差矩阵的极限特征值分布并简化估计协方差矩阵的特征值分布, 从而实现低功率信号检测。仿真结果表明, 与同类算法相比新算法检测精度高, 所需功率低, 具有较好的理论及实用价值。
关键词: 无线通信; 信号检测; 自由概率理论; 随机矩阵; 极限特征值分布

A New Low Power Signal Detection Algorithm

Wang Lei Zheng Bao yu Cui Jing wu

(Institute of Signal Processing and Transmission, Nanjing University of Posts and Telecomm.,
Jiangshu, Nanjing 210003, China)

Abstract: Aiming at green communication in the future, how Free Probability Theory (FPT) can be used to aid in low power signal detection for communication systems was studied in this paper. FPT is an invaluable tool for describing the asymptotic behavior of many systems. Sample covariance matrices of communication systems with noise were the starting point in our FPT-based scheme. Multiplicative free deconvolution was shown to be a method which can aid in expressing limit eigenvalue distributions for sample covariance matrices, and to simplify estimators for eigenvalue distributions of covariance matrices. Then low power signal detection was carried out. Simulation results show new scheme is more precise than congeneric method with low power and with theoretical and practical value.
Keywords: Wireless communication; Signal detection; Free Probability Theory; Random Matrices; Limiting eigenvalue distribution

1 引言

随机矩阵理论 (RMT), 尤其是随机样本协方差矩阵的极限特征分布在刻画大维数系统时是一种非常有用的数学工具, 在数字通信、核物理、金融数学等方面均有成功的应用^[1]。
无线通信系统中一个典型的随机矩阵模型为信号-噪声模型, 其样本协方差矩阵可表示为:

基金项目: 国家自然科学基金资助项目 (60972039); 江苏省自然科学基金资助项目 (BK2007729); 江苏省高校自然科学基金资助项目 (06KJA51001); 南京邮电大学攀登计划资助项目 (NY207061)

$$W_n = \frac{1}{N} (R_n + \sigma X_n)(R_n + \sigma X_n)^H \quad (1)$$

其中 R_n 和 X_n 是 $n \times N$ 相互独立的随机矩阵, X_n 中的元素是独立同分布的标准 (均值为 0 方差为 1) 复高斯随机变量。式(1)可以看成是随机向量 $r_n + \sigma x_n$ 的样本协方差矩阵。 r_n 可以认为是包含系统特征 (例如雷达中的到达方向或信道估计中的冲击响应) 的向量。 x_n 表示加性噪声, σ 是表示噪声强度的参数。在本文中 n 和 N 满足

$$\lim_{n \rightarrow \infty} \frac{n}{N} = c \quad (2)$$

即观测值数和系统参数的个数具有相同的增长速率。这是许多信号处理应用中都具备的条件, 这是在信号特征不变的区间内我们只能获得有限的观测值。

促使我们研究这类问题的是如下情景: 假定 N 个观测值由 n 个传感器获得, 每个传感器的观测值都是在信号来源和数量均未知的情况下获得。此外, 每个传感器都受到噪声的影响。这些传感器构成了一个随机向量 $r_n + \sigma x_n$, 可由观测值构成样本协方差矩阵 W_n 。若 W_n 已知, 我们感兴趣的是如何由此得到随机向量 r_n , 从而进一步分析系统(1)。在这种情况下, 可以在下面这些量之间建立联系:

- A. W_n 的特征值分布,
- B. $\Gamma_n = \frac{1}{N} R_n R_n^H$ 的特征值分布,
- C. 协方差矩阵 $\Theta_n = E(r_n r_n^H)$ 的特征值分布。

在文献[2]中, Dozier 和 Silverstein 给出了通过解一个给定的方程由 B 来估计 A 。但没有提供解决这个问题的算法。实际上, 许多应用中想获取系统信息时是要求由 A 来估计 B , 遗憾的是文献[2]没有解决这个问题。最近, 在文献[3]中, 我们认识到文献[2]中的框架实际上阐明的是“乘法自由卷积(*multiplicative free convolution*)”的概念。而且文献[3]介绍了“自由解卷积(*free convolution*)”的概念, 并类似于通过 B 来估计 A 那样提出了通过 A 估计 B 的方法。

C 可以有由 G_2 -估计器^[4]获得, 它给出了一个协方差矩阵 Stieltjes 变换的一致估计, 估计的基础是样本协方差矩阵的 Stieltjes 变换。 G -估计器已经在很多应用中非常有效, 却仍缺少对它的原理的直观的解释。在文献[3]中, 作者还指出 G_2 -估计器也可以在乘法自由卷积理论框架下获得。这就为找到 B 提供了计算方法。注意到, 像文献[5]中证明的那样, C 可以直接得到而不必通过 B 。但是, 后者没有提供一个统一的框架来计算全部的特征值分布, 而只是给出了一些矩。

除上述数学框架之外, 文献[6]还研究了“自由解卷积(*free de-convolution*)”问题。有趣的是, 乘法自由解卷积操作起来可以很方便, 该文详细描述和证明这个问题。它与信号处理中经常涉及到的问题有着紧密联系。将涉及到通信系统, 信号子空间的秩的估计, 噪声方差估计及信道容量估计等问题。

我们研究如何运用 FPT 进行无线通信系统中的低功率信号检测。文章第二部分给出自由概率理论的相关基本概念, 包括乘法和加法自由卷积和解卷积。第三部分将具体介绍如何运用 FPT 进行无线通信系统的低功率信号检测, 并通过仿真对算法性能进行分析和比较。

2 自由卷积理论

上世纪 80 年代, 在 Voiculescu 的努力下^[7-10], 自由概率(*free probability*)^[11]理论已发展成

为一个较完整的研究领域。自由概率理论中的一些基本概念是比较抽象的，其目的是引入一个类似于经典概率理论中“独立”的概念，使之能够适用与非交换随机变量（如矩阵）。这些更一般的随机变量作为“非交换概率空间(non-commutative probability space)”中的元素可以通过一个序对 (A, ϕ) 来定义，其中 A 是具有幺元 I 的 $*$ -代数系统， ϕ 是 A 上一个规格化 (i.e. $\phi(I)=1$) 的线性函数。 A 中的元素称为随机变量。本文中 A 由 $n \times n$ 矩阵或随机矩阵组成。对矩阵来说， ϕ 是标准化的迹 tr_n ，定义为（对于 $\forall a \in A$ ）

$$\text{tr}_n(a) = \frac{1}{n} \text{Tr}(a) = \frac{1}{n} \sum_{i=1}^n a_{ii} \quad (3)$$

对于随机矩阵时， ϕ 是线性函数 τ_n ，定义为

$$\tau_n(a) = \frac{1}{n} \sum_{i=1}^n E(a_{ii}) = E(\text{tr}_n(a)) \quad (4)$$

这类 $*$ -代数中的幺元是 $n \times n$ 的单位矩阵 I_n 。这类非交换概率空间都具有迹性，即 ϕ 满足迹的性质 $\phi(ab) = \phi(ba)$ 。其中类似于经典概率论中“独立”的概念称作“自由(freeness)”。

定义 1: 一族幺元 $*$ -子代数 $(A_i)_{i \in I}$ 系统被称为一个自由族(free family)，如果下式成立

$$\left\{ \begin{array}{l} a_j \in A_{i_j} \\ i_1 \neq i_2, i_2 \neq i_3, \dots, i_{n-1} \neq i_n \\ \phi(a_1) = \phi(a_2) = \dots = \phi(a_n) = 0 \end{array} \right\} \Rightarrow \phi(a_1 \cdots a_n) = 0 \quad (5)$$

随机变量族 a_i 被称为自由族，如果在它们上面定义的代数系统是由一个自由族生成的。

定义 2: 称一个概率空间 (A_n, ϕ_n) 中的随机变量序列依分布收敛，如果对于任意 $m_1, \dots, m_r \in \mathbb{Z}$ ， $k_1, \dots, k_r \in \{1, 2, \dots\}$ ，在 $n \rightarrow \infty$ 时，极限 $\phi_n(a_{nk_1}^{m_1} \cdots a_{nk_r}^{m_r})$ 存在。如果这些极限对于非交换概率空间 (A, ϕ) 和自由变量 $a_1, a_2, \dots \in (A, \phi)$ 可以写成 $\phi(a_{nk_1}^{m_1} \cdots a_{nk_r}^{m_r})$ ，我们就称 a_{n1}, a_{n2}, \dots “渐近自由(asymptotically free)”。

渐近自由是一个非常有用的概念，因为很多类型的矩阵在维数趋于无穷时都渐近自由。例如，随机矩阵 $\frac{1}{\sqrt{n}} A_{n1}, \frac{1}{\sqrt{n}} A_{n2}, \dots$ ，这里 A_{ni} 是 $n \times n$ 维的，其元素满足独立标准高斯分布。这就是文献[11]中给出的 $\frac{1}{\sqrt{n}} A_{ni}$ 渐近自由。这种情况下 $\frac{1}{\sqrt{n}} A_{ni}$ 的极限分布被称为“圆律(circular law)”：其特征值均匀地分布到复平面上的单位圆内。

公式(5)使我们可以计算自由变量 a_1 和 a_2 的任何阶混合矩。当可以用矩序列唯一地把概率测度（紧支撑概率测度）统一起来时， $a_1 + a_2$ 和 $a_1 a_2$ 的分布成为两个新的概率测度，它们是仅与 a_1, a_2 的矩相关的概率测度。因此我们可以在概率测度集合上定义两个新算子：对于两个自由随机变量的和，定义加法自由卷积(Additive free convolution)

$$\mu_1 + \mu_2 \quad (6)$$

对于两个自由随机变量的积，定义乘法自由卷积(Multiplicative free convolution)

$$\mu_1 \times \mu_2 \quad (7)$$

这些新算子可以用来估计自由随机变量的和或积的谱。例如，若 a_{1n} 的渐近特征分布逼近于 μ_1 ， a_{2n} 的渐近特征分布逼近于 μ_2 ，我们有 $a_{1n} + a_{2n}$ 有渐近特征分布逼近于(6)，所以可以用它来估计大维矩阵的特征值。通常情况下对矩阵组合的特征值估计是困难的，除非我们对特征向量的结构做出某些假设。使随机矩阵适于自由概率理论（使随机矩阵自由）的假设为：具有一致分布的特征向量结构（即，在某种意义上特征向量点在所有方向具有相同的概率）。

加法和乘法自由解卷积(additive and multiplicative free de-convolution)的概念也非常有用，

具体如下：

定义 3: 给定概率测度 μ 和 μ_2 。当存在一个唯一的概率测度 μ_1 分别满足

$$\mu = \mu_1 + \mu_2, \quad \mu = \mu_1 \times \mu_2$$

可以写成

$$\mu_1 = \mu - \mu_2, \quad \mu_1 = \mu / \mu_2$$

我们说 μ_1 是 μ 和 μ_2 “加法自由解卷积”和“乘法自由解卷积”。

许多情况下一些概率测度是作为大维随机矩阵的极限存在的。其中一个重要的测度是 Marchenko Pastur 律 μ_c ^[1]，在自由概率理论中也称为“自由泊松分布”。它的概率密度函数为

$$f^{\mu_c}(x) = (1 - \frac{1}{c})^+ \delta(x) + \frac{\sqrt{(x-a)^+(b-x)^+}}{2\pi cx} \tag{8}$$

这里 $(z)^+ = \max(0, z)$ ， $a = (1 - \sqrt{c})^2$ ， $b = (1 + \sqrt{c})^2$ 。图 1 中给出了 μ_c 在 c 取一些值时的曲线。 μ_c 描述的是 Wishart 矩阵渐进特征值分布。Wishart 矩阵形式为 $\frac{1}{N}RR^H$ ，这里 R 是 $n \times N$ 随机矩阵，其元素满足独立标准高斯分布。 μ_c 为 $\frac{n}{N} \rightarrow c$ 当 $n \rightarrow \infty$ 时的极限分布。

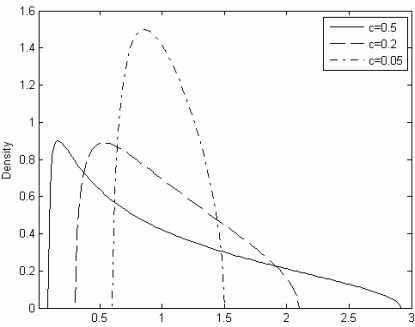


图 1 不同的 Marchenko Pastur 律 μ_c

Marchenko Pastur 律在非高斯情况下也是成立的。

通常用测度的概率密度来描述自由卷积，因为它提供了特征值分布。这方面一个非常重要的工具是 Stieltjes 变换 (Stieltjes transform)^[1]。对于概率测度 μ ，这是一个 $C^+ = \{z \in C : \text{Im } z > 0\}$ 上的解析函数，定义为：

$$m_\mu(z) = \int \frac{1}{\lambda - z} dF^\mu(\lambda) \tag{9}$$

这里 F^μ 是 μ 的累计分布函数。所有的 μ 被认为在实轴的正半轴有支撑 (support)。这样的 μ ， m_μ 可以被解析延续到实轴的负半轴，这时 m_μ 的值是实数。如果 μ 具有紧支撑 (compact support)， $m_\mu(z)$ 可以用劳伦级数 (Laurent series) 展开，展开的系数是 μ 的 k 阶矩 μ_k ：

$$m_\mu(z) = -\frac{1}{z} \sum_{k=0}^{\infty} \frac{\mu_k}{z^k} \tag{10}$$

一个使用起来方便的 Stieltjes 变换的逆公式为：

$$f^\mu(\lambda) = \lim_{\varpi \rightarrow 0^+} \frac{1}{\pi} \text{Im} [m_\mu(\lambda + j\varpi)] \tag{11}$$

3 FPT在无线信号检测中的应用

在无线通信系统中,常常需要检测 OFDM 系统中接收到信号的功率。用 N 表示子载波数, L 表示循环前缀长度,用 h 表示 $L+1$ 径频率选择性信道,方差为 σ_h^2 , 在 OFDM 系统中接收机收到的信号向量为:

$$y_i = F(TP^{\frac{1}{2}}s_i + n_i) \quad (12)$$

这里 T 是 $N \times (N+L)$ Toeplitz 信道矩阵:

$$T(h) = \begin{pmatrix} h_L & \dots & h_0 & 0 & \dots & 0 \\ 0 & \ddots & & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & & \ddots & 0 \\ 0 & \dots & 0 & h_L & \dots & h_0 \end{pmatrix} \quad (13)$$

F 是 FFT 矩阵 $[F]_{kl} = \exp(-2\pi j \frac{kl}{N})$, y_i 为 $N \times 1$ 接收向量, P 是对角功率矩阵, s_i 所传输信息, n_i 是零均值加性高斯白噪声, 方差为 σ^2 。

通常确定功率的方法是当信道矩阵和噪声已知情况下找到 y_i 协方差矩阵的特征值。即

$$\Theta = E(y_i y_i^H) = \sigma_h^2 P + \sigma^2 I \quad (14)$$

但实际中只能获得协方差矩阵的估计而不知道用户的扩频序列, 只可以假定噪声方差已知。实际上, 通过下式计算样本协方差矩阵:

$$\hat{\Theta} = \frac{1}{L} \sum_{i=1}^L y_i y_i^H \quad (15)$$

这类传统方法在 L 和 n 在相近时是非常不充分的。

自由解卷积理论框架非常适合这种情况并可以在不需要知道扩频序列的情况下确定用户功率。实际上, 根据 Wishart 分布的性质, 样本协方差矩阵与统计协方差矩阵有如下联系:

$$\hat{\Theta} = \Theta^{\frac{1}{2}} X X^H \Theta^{\frac{1}{2}} \quad (16)$$

这里协方差矩阵为

$$\Theta = E(y_i y_i^H) = \sigma_h^2 P + \sigma^2 I \quad (17)$$

X 是 $N \times L$ 独立同分布零均值高斯分布矩阵。

联合(16)和(17)并注意到 $\frac{1}{L} X X^H$ 是 Wishart 矩阵且它们的分布逼近 $\mu_{\frac{N}{L}}$, 可以得到下面的渐进自由方程

$$\left(\mu_{\sigma_h^2 P} \boxplus \mu_{\sigma^2 I} \right) \boxtimes \mu_{\frac{N}{L}} = \mu_{\hat{\Theta}} \quad (18)$$

如果知道噪声方差, 通过加法和乘法自由解卷积, 可以得到:

$$\mu_{\sigma_h^2 P} = \left(\mu_{\hat{\Theta}} \boxtimes \mu_{\frac{N}{L}} \right) \boxminus \mu_{\sigma^2 I} \quad (19)$$

在接下来的仿真中, 我们使用了子载波数 $N=64$, 循环前缀数 $L=16$, 噪声方差 $\sigma^2=1$ 。求解自由(解)卷积时可以采用“矩-累积量法”^[6]。功率矩阵 P 为对角阵, 等概率分成三部分, 取值分别为, 0, 0.5, 和 1, 即

$$\mu_P = \frac{1}{3} \delta_0 + \frac{1}{3} \delta_{0.5} + \frac{1}{3} \delta_1 \quad (20)$$

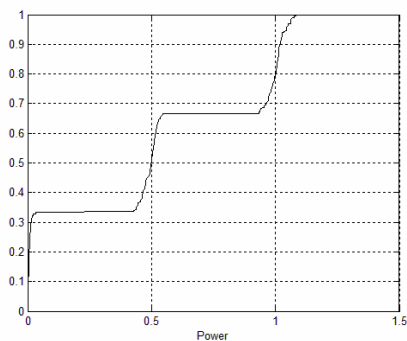


图 2 功率累积分布函数

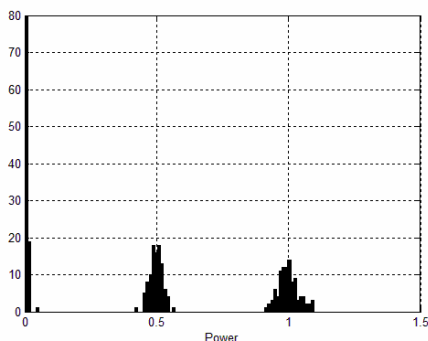


图 3 特征值频率直方图

在图 2 中，通过解方程(18)来估计 μ_p 的累积分布函数，采用“矩-累积量法”计算前三阶矩。并使用这些矩通过 Newton-Girard 公式来估计特征值。对每个观测值数执行 100 次，并用这 100 次特征值的平均值来估算累积分布函数，当然也可以提高矩的阶数而减少执行的次数。如图 2 所示，得到接近式(20)的累积分布函数。图 3 为相应的特征值频率直方图。仿真结果还说明了基于 FPT 的方法收敛速度快，在有限维下仍然有效。

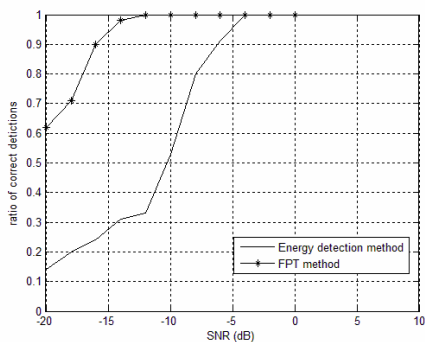


图 4 与能量检测方法的性能比较

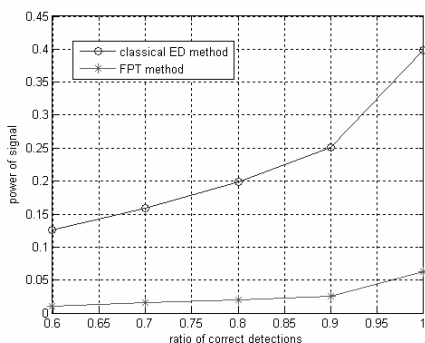


图 5 达到相同检测正确率是所需功率比较

图 4 给出了基于 FPT 的算法与传统能量检测方法的性能比较。由图可见，基于 FPT 的检测算法在相同信噪比下性能明显优于传统能量检测算法，尤其是在低信噪比情况下。图 5 给出了在背景噪声相同的情况下 ($\sigma^2 = 1$)，达到相同检测正确率的时两种算法所需的信号功率的对比情况。可见，基于 FPT 的算法达到任一正确率时所需要的信号功率明显低于传统能量检测算法，从而更接近未来的低功率绿色通信。

4 结论

本文研究了无线通信系统中基于自由概率理论的低功率信号检测算法。文中分析了算法的基本原理，并通过仿真分析比较了算法性能。仿真结果表明基于 FPT 的算法收敛速度快，在有限维下仍非常有效，同时还表明新算法可以明显降低检测信号时需要的信号功率，适合于解决无线通信系统中的低功率信号检测问题。

参 考 文 献

- [1] TULINO A M, VERDU S. Random Matrix Theory and Wireless Communications [M]. Hanover, USA: Now Publishers Inc., 2004.
- [2] DOZIER B, SILVERSTEIN J. On the empirical distribution of eigenvalues of large dimensional information-plus-noise type matrices [J]. Journal of Multivariate Analysis, vol. 98, Issue 4, pp 678-694, 1 April 2007.
- [3] RYAN Ø. Multiplicative free convolution and information-plus-noise type matrices [J]. Annals of Applied Probability, 2007.
- [4] GIRKO V L. Ten years of general statistical analysis [J]. <http://generalstatistical-analysis.girko.freewebspace.com/chapter14.pdf>, 2007.
- [5] RAO N, EDELMAN A. Free probability, sample covariance matrices and signal processing [C]. ICASSP, Toulouse, July 2006, (5):V-V
- [6] RYAN Ø, DEBBAH M. Free deconvolution for signal processing applications [J]. IEEE Trans. On IT, 2007, 11(1): 569-571.
- [7] VOICULESCU D. Addition of certain non- commuting random variables [J]. J. Funct. Anal., 1986, (66) :323-335.

- [8] VOICULESCU D. Multiplication of certain non-commuting random variables [J]. J. Operator Theory, 1987, 2(18): 223-235.
- [9] VOICULESCU D. Circular and semicircular systems and free product factors [J]. Inv. Math., vol. 92, 1990.
- [10] VOICULESCU D. Limit laws for random matrices and free products [J]. Inv. Math., 1991, (104): 201-220.
- [11] HIAI F, PETZ D. The Semicircle Law, Free Random Variables and Entropy [M]. American Mathematical Society, Providence, RI, 2006.

作者简介:

王 磊 (1977-), 男, 黑龙江人, 博士生, 研究方向为现代通信中的职能信号与信息处理。

郑宝玉 (1945-), 男, 福建闽侯人, 教授、博士生导师; 研究方向为智能信号与信息处理等。

崔景伍 (1955-), 女, 河北里县人, 高级工程师, 研究方向为无线通信与网络信号处理等。

宽带雷达信号接收波束形成的方法

章华璠 王盛利

(南京电子技术研究所, 江苏南京 210013)

摘要: 常规波束形成方法具有运算量低、稳健性高、不需要目标先验知识等优点, 在雷达系统中得到广泛应用。但这些应用都是基于窄带雷达信号。目前, 对于宽带雷达信号波束形成方法的研究还比较少。针对宽带雷达信号, 文中提出了基于聚焦变换的波束形成方法, 将宽带数字阵列接收信号聚焦到一个窄带上进行数字波束形成处理。通过聚焦变化的方法, 使宽带数字波束形成与窄带数字波束形成一样简单, 有效, 同时克服了直接采用窄带数字波束形成方法的缺陷。通过仿真实验, 对该方法进行了验证。结果表明, 该方法能获得很好的波束形成性能, 使接收信号波形基本不失真。

关键词: 阵列信号处理; 宽带信号波束形成; 聚焦变换; 窄带分解

The Method of Receiving Beam-forming for Wideband Radar Signal

Zhang Hua luan Wang Sheng li

(Nanjing Research Institute of Electronic Technology, Nanjing 210013 , China)

Abstract: The method of conventional beam-forming is widely used in radar system because of its low computation quantity、high robustness and needlessness of transcendental knowledge. But all this applications are based on narrowband radar signal. Currently, the digital beam-forming based on wideband radar signal is still rare. This paper presents a method of wideband radar signal beam-forming, which is based on focusing transform and processes the wideband signal by focusing the wideband signal on a narrowband case. The approach makes the method for wideband signal beam-forming as easy as narrowband signal beam-forming and also resolves the shortcomings caused by applying the narrowband signal beam-forming method directly for the wideband case. The performance of the method is tested by the computer simulation. The simulation results show that the performance of wideband signal beam-forming is good and the waveform of the receiving signal is distortionless.

Keywords: array signal processing; wideband signal beam-forming ; focusing transform; narrowband decomposition

0 引言

雷达性能的好坏, 在很大程度上依赖于天线阵列接收波束的形成的能力。例如, 波束的精确指向, 干扰对消, 旁瓣抑制等能使雷达的接收信号不失真, 对信号的后续处理至关重要。

波束形成的目的是使信号到达各阵元的延时引起的相位变化得到补偿,以使接收期望方向的信号达到最大。由延迟引起的信号的相位变化跟信号的波长 λ 或频率 f 有关。在窄带的情况下,由于带宽相对与载波频率非常小(带宽与载波频率的比值小于1%),所以在波束形成时可以忽略频率的变化对方向矢量的影响,用单一的权向量就可形成所需的波束。

但是宽带信号情况就不同了,宽带信号由于方向矢量随信号频率的变化影响较大(带宽与载波频率的比值大于10%),使波束形成不能像在窄带信号中处理一样,使用单一的权向量。对宽带信号处理,一般的做法是先利用离散傅里叶变换进行宽带信号的窄带分解,再对每一个窄子带进行处理,然后综合各个处理结果。所以该方法能将窄带信号波束形成的各种方法直接应用到宽带信号波束形成中,即将宽带信号的每个窄子带进行所需的波束形成方法,再综合各子带的结果。为了提高性能,减少计算量,本文采用了由H.Wang提出的相干信号子空间算法[1]。其主要用来计算空间信号的功率谱估计,尤其在声学领域的应用特别常见[3][4]。其基本思想就是把频带内不重叠的频率点通过聚焦变换到参考频率点上,聚焦后得到单一频率点的信号,再应用窄带波束形成方法进行波束形成。

1 宽带信号模型

假设有 N 个带宽为 B 的远场宽带信号作用于宽带数字雷达阵列天线上,其中天线阵列由 M 个阵元组成的均匀线阵,这里假设阵元数等于通道数。噪声为通道一致加性噪声,信号与噪声,各阵元间的噪声都不相关。如图1所示:

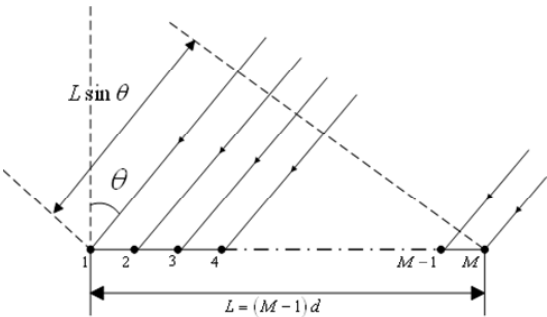


图1 阵列模型

则第 m 个阵元接收的数据(不考虑增益时)为:

$$x_m = \sum_{n=1}^N s[t - \tau_{mn}] + n_m(t), \quad m = 1, 2, \dots, M \tag{1}$$

其中 τ_{mn} 是第 n 个信号相对阵列参考点的时间延时。然后对每个通道的接收数据进行 J 点DFT,就可以得到如下的宽带信号模型:

$$X(f_j) = A(f_j)S(f_j) + N(f_j), \quad j = 1, 2, \dots, J \tag{2}$$

式中 $X(f_j)$, $S(f_j)$, $N(f_j)$ 分别为对应于频率 f_j 的接收数据,信号及噪声的DFT变换,表达式如下:

$$X(f_j) = [X_1(f_j) \quad X_2(f_j) \quad \dots \quad X_M(f_j)]^T \tag{3}$$

$$S(f_j) = [S_1(f_j) \quad S_2(f_j) \quad \cdots \quad S_N(f_j)]^T \quad (4)$$

$$N(f_j) = [N_1(f_j) \quad N_2(f_j) \quad \cdots \quad N_M(f_j)]^T \quad (5)$$

$$A(f_j) = [a_1(f_j) \quad a_2(f_j) \quad \cdots \quad a_N(f_j)] \quad (6)$$

$$\begin{aligned} a_n(f_j) &= [a_{1n}(f_j) \quad a_{2n}(f_j) \quad \cdots \quad a_{Mn}(f_j)]^T \\ &= [e^{-j2\pi f \tau_{1n}} \quad e^{-j2\pi f \tau_{2n}} \quad \cdots \quad e^{-j2\pi f \tau_{Mn}}]^T \end{aligned} \quad (7)$$

需要强调的是，式(2)中的 J 时指将带宽为 B 的信号划分为 J 个子带，所以对应于不同频率 f_j ，有 J 个跟式(2)一样的等式。

2 窄带波束形成方法

为了分析方便，我们假设只有一个远场信号源作用于阵列。设波束形成的加权矢量为：

$$W = [w_1 \quad w_2 \quad \cdots \quad w_M]^T \quad (8)$$

式中， W 的值是根据所采用的波束形成方法得到的加权值。根据窄带的方法直接用一组加权值，则波束形成的输出是：

$$\begin{aligned} Y(f_j) &= W^T X(f_j) \\ &= W^T a(f_j) S(f_j) + W^T N(f_j) \end{aligned} \quad (9)$$

则基于窄带方法的宽带信号的波束形成输出为：

$$Y(f) = \sum_{j=1}^J Y(f_j) \quad (10)$$

对上式进行 IFFT 就得到了宽带信号波束形成的时域输出。上述的方法就是基于窄带方法的宽带信号的波束形成。

根据该方法进行仿真，信号 $x(t)$ 采用带宽 $B=400\text{MHz}$ ，载波频率 $f_c=1\text{GHz}$ 的线性调频信号，采样频率 $f_s=1\text{GHz}$ 。 $x(t)$ 的时域波形如图 2 所示：

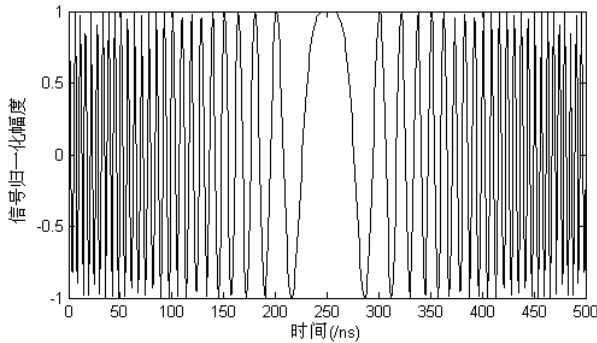


图 2 发送信号时域波形

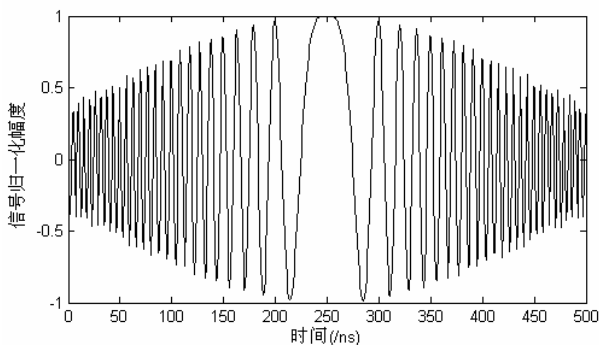


图 3 波束形成输出

从图 3 可以看出，直接用窄带的方法（即只用一组加权值）进行波束形成，输出波形严重失真，特别是信号的高频部分，幅度最大的损失量为原来的 33.3%。在这种情况下，必将对信号的后续处理带来严重的后果。

3 宽带波束形成方法

由上述可知，宽带信号波束形成不能用窄带信号的方法，直接用一组加权值进行波束形成。这是因为由延迟引起的信号的相位变化跟信号的波长 λ 或频率 f 有关。窄带情况下，信号的带宽相对载波频率比较小，所以信号的波长 λ 或频率 f 的变化可以忽略不计。宽带情况下，信号的带宽大于载波频率的 10%，如果忽略波长 λ 或频率 f 的变化，直接用窄带的方法就会得到图 2 所示的结果。

根据上面的分析，采用频域方法对宽带信号进行波束形成，原理图如图 4 所示：

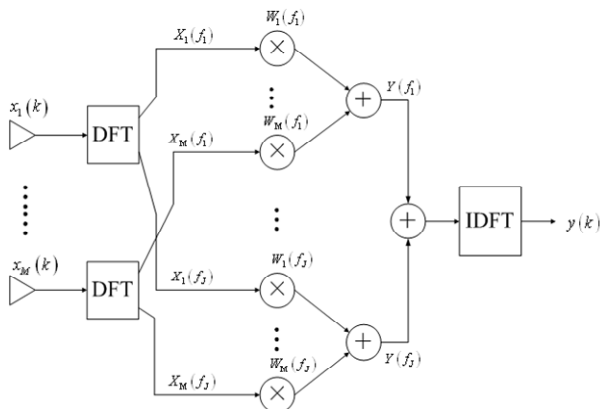


图 4 波束形成原理图

图中， $x_m(k), (m=1, 2, \dots, M)$ 表示第 m 号阵元接收到时序为 k 的信号； $X_m(\omega_j), (j=1, 2, \dots, J)$ 为第 m 号阵元接收信号进行 J 点 FFT 后，数字频率为 j 的宽带信号分量； $W_m(\omega_j)$ 为第 m 号阵元上对数字频率为 j 的信号分量的加权值； $Y(\omega_j)$ 为数字频率为 j 信号分量的波束形成输出； $y(k)$ 为宽带信号波束形成的最终输出。

该方法首先把每个阵元上接收到的信号通过 DFT 转换到频域，然后对宽带信号的每一个频率分量施加不同的加权值得到此频率分量上的波束输出：

$$Y(f_j) = W(f_j)^T X(f_j) \quad (11)$$

再综合各频率分量的输出，得到：

$$Y(f) = \sum_{j=1}^J Y(f_j) \quad (12)$$

通过 IDFT 就可以得到宽带信号的时域输出。由于做 J 点 DFT 相当于把宽带信号在频域上划分为 J 个子带，并对每一个窄子带进行窄带波束形成。所以窄带波束形成的方法都能应用到宽带信号的波束形成中。

4 基于聚焦变换的波束形成方法

由式(11)可知，上述方法需要 J 组不同的加权值。基于聚焦变换的方法只要求一组加权值。它的基本原理是把频带内不重叠的频率点通过聚焦变换到参考频率点上，聚焦后得到单一频率点的信号，这样只需求聚焦频率点的加权值。聚焦变换的方法主要集中在聚焦矩阵的求解上。按照文献[2]中聚焦的概念，针对约束方向的聚焦模型如下：

$$T(f_j) a(f_j) = a(f_0) \quad (13)$$

式中，下标 0 表示对应于聚焦的参考频率点。文献[3]对聚焦参考频率点的选择设计了下述最优化的问题：

$$\begin{cases} \min_j \min_{T_j} \sum_{j=1}^J W_j \|a_0 - T_j a_j\|^2 \\ s.t. \quad T_j^H T_j = T_j T_j^H = I, \quad j=1, 2, \dots, J \end{cases} \quad (14)$$

根据文献[3]分析结果，最优聚焦参考频率点在中心频率的附近。

其中：

$$T(f_j) = \begin{bmatrix} \frac{a_1(f_0)}{a_1(f_j)} & & \\ & \ddots & \\ & & \frac{a_M(f_0)}{a_M(f_j)} \end{bmatrix} \quad (15)$$

所以根据上式，对原理图 4 进行改进，得到如下聚焦变换方法的原理图：

图 5 中，对输入信号的 $X(f_j)$ 进行预处理得到：

$$\begin{aligned} X'(f_j) &= T(f_j) X(f_j) \\ &= T(f_j) [a(f_j) S(f_j) + N(f_j)] \\ &= a(f_0) S(f_j) + T(f_j) N(f_j) \end{aligned} \quad (16)$$

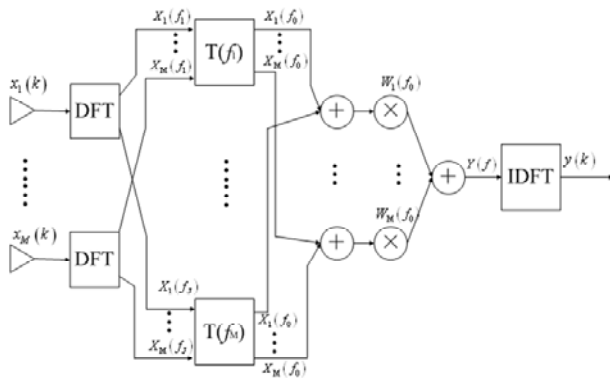


图 5 基于聚焦变换的波束形成原理图

则各窄带分量的方向矢量都聚焦到了参考频率点上。所以波束形成时可以用单一的权向量 $W(f_0)$ 。则(12)可变换为如下形式：

$$\begin{aligned}
 Y(f) &= W(f_0)^T \sum_{j=1}^J X'(f_j) \\
 &= W(f_0)^T \sum_{j=1}^J T(f_j) X(f_j)
 \end{aligned} \tag{17}$$

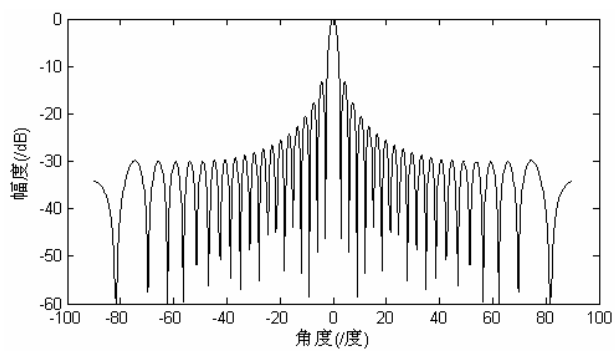
上式即为基于聚焦变换的宽带信号波束形成的算法。与(11)式相比，式(17)对 W 求取只要进行一次就可以了，而(11)要进行 J 次(J 越大波束形成的效果越好)。根据波束形成方法的不同， W 的求取算法也不同，相对于式(17)，式(11)求取 W 的计算量要高的多。所以用聚焦的方法，可以减少波束形成的计算量。

5 仿真结果

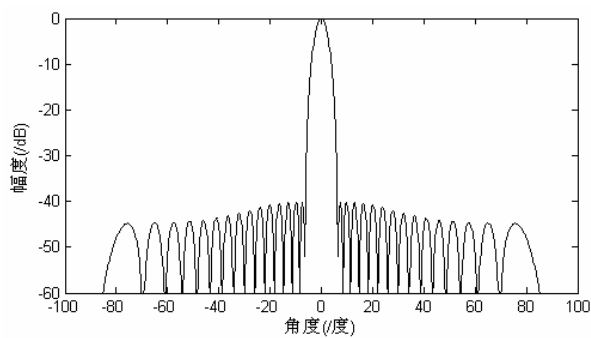
考虑一个由 32 个阵元组成线阵，阵元间距为信号中心频率对应的半波长。源信号为线性调频信号，中心频率为 1GHz，带宽为 0.4GHz。噪声为零均值的高斯过程与信号有相同的带宽。阵列接收信号通过 FFT 分解为 129 个独立的子带。本文将中心频率 f_0 作为参考频率。

图 6 中，(a)图是宽带信号通过窄带分解并进行聚焦变换得到的波束方向图，其第一副瓣值为 13.47dB；(b)图是采用 32 阶，第一副瓣值为 -40dB 泰勒窗加权后得到的波束方向图；可以看出，经过泰勒加权后，方向图副瓣小于 -40dB，但主瓣展宽了。

图 7 中，(a)图为宽带接收信号未处理的时域波形；(b)图为经过聚焦变换波束形成方法得到的接收信号时域波形图。很明显，通过聚焦变换的方法能恢复信号的波形。

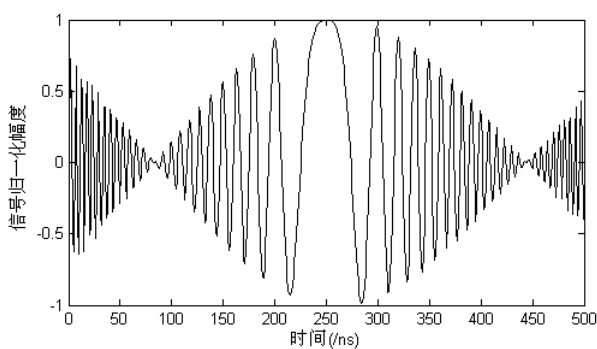


(a)

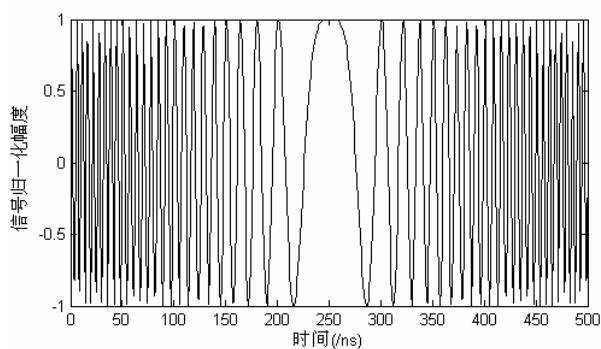


(b)

图 6 波束指向图



(a)



(b)

图 7 信号时域波形图

6 结束语

本文分析了宽带信号波束形成与窄带信号的差异,根据宽带信号的特点对其进行窄带分解,并用聚焦变换算法对宽带信号进行数字波束形成。通过仿真实验可以看出,该方法能很好的对宽带信号进行波束形成。

参 考 文 献

- [1] H.WANG and M .KAVEH. Coherent signal-subspace processing for the detection and estimation of Allgles of multiple wide-band sources[J]. IEEE Trans. on ASSP,Vol.33,No.4, pp: 823-831, Aug,1985.
- [2] Valaee.S and KABAL.P. The optimal focusing subspace for coherent signal subspace processing[J].IEEE.Trans,SP,1996,44(3):752—756
- [3] 陈建峰,黄建国,张群飞.一种宽容的宽带加权聚焦定向算法[J].声学学报,2000,25(1):16—20
- [4] 张德明,郭良浩,张仁和.基于聚焦变换的宽带信号方位稳健估计[J].声学学报,2005,Vol.30 NO.4
- [5] H.COX, R .M .ZESKIND and M .M .OWEN, “Robust adaptive beamforming[J], IEEE Trans. On AP, Vol.35, Oct,1987.

作者简介:

章华鑫,1983年生,男,硕士研究生。研究方向为雷达信号处理。

王盛利,1957年生,男,研究员级高级工程师,工学博士。研究方向为雷达系统设计以及雷达信号处理。

基于瞬时距离-瞬时多普勒的高速机动目标成像

句彦伟 李士国 于 立

(南京电子技术研究所, 江苏 南京 210039)

摘 要: 传统 RD 算法 (Range Doppler algorithm) 的 ISAR 成像 (Inverse synthetic aperture radar imaging) 是基于目标匀速运动的假设, 对于高速且机动运动的目标, 高速性会使目标一维距离像展宽模糊; 目标机动性会使得利用 RD 算法散射点展宽, 二维图像模糊。通过深入地分析, 文章给出一种利用短时傅立叶变换距离向压缩解决目标高速性、通过时频分析解决目标机动性的 ISAR 成像方法, 仿真试验验证了所提出的成像算法的有效性和潜力。
关键词: RD 算法; 短时傅立叶变换; 时频分析; ISAR 成像

Imaging of Fast-moving and Maneuvering Target Based on Instantaneous Range and Instantaneous Doppler Technique

Ju Yan wei Li Shi guo Yu Li

(Nanjing Research Institute of Electronics Technology, Nanjing 210039, China)

Abstract: The traditional range Doppler algorithm is suitable to the stationary target. When the fast target moves irregularly, the scatter points of target become wide and ISAR image blurry. First, this paper analyses the imaging principle of fast and maneuvering target; then, we take advantage of short-time Fourier transform to resolve the effect of fast speed, and use time-frequency analysis technique to obtain good instantaneous image of the target. Finally, the simulated data demonstrate the validity of the method proposed.
Keywords: RD algorithm; short-time Fourier transform; time-frequency analysis; ISAR imaging

1 引言

逆合成孔径雷达(ISAR)通过雷达发射大带宽信号获得径向距离分辨力, 通过目标与雷达相对转动获得横向分辨力。ISAR 成像在战略防御、反导、目标识别、高精度测距、战术武器等方面都有重要的应用价值^{[1]-[3]}。随着 ISAR 分辨率的提高和被探测目标范围的扩展, 近年来出现了对空间轨道目标和弹道目标进行高分辨成像的需求, 这类目标的运动速度很高, 它们在脉冲持续期间的径向运动距离与距离分辨单元大小可比拟, 因此解调混频输出不再是单频分量, 而是调频斜率相同的多分量线形调频信号, 此时在频域获得的距离像将产生畸变, 进而引起 ISAR 二维像的模糊。另外, 实际雷达与目标相对径向速度并不是常数, 不可能平稳飞行, 而是在做不同程度的机动飞行, 这样运动过程就很复杂, 这时各个散射点的多普勒频率是时变的, 用傅里叶变换进行谱分析得到目标图像会变得模糊。

针对目标高速问题,国内外学者提出了一些速度补偿方法,其中文献[4]提到使用 RAT, RELAX 方法估计混频输出信号的调频率,文献[5]提出使用谱包络最小熵准则估计线性调频信号参数。这些方法的共同点是:通过估计混频输出信号的调频斜率获得目标径向运动速度估计值,然后用该估计值构造补偿信号对输出信号进行补偿,以消除高速运动带来的影响,补偿后的信号等效为目标低速运动时的混频输出,此时使用经典解线调处理就能够实现目标高分辨成像。不同于以上方法,我们采用瞬时距离-多普勒方法来解决目标高速性,即用短时傅里叶变换代替传统的傅里叶变换进行距离压缩得到目标一维距离像^[6]。而对于目标机动性,我们提出使用 Gabor 时频分析瞬时成像方法代替 RD 成像,分析信号不同时刻的频率,从而得到目标的瞬时距离-瞬时多普勒像^{[7]-[8]}。利用时频分析方法不仅可以得到清晰的目标图像,而且不需要对各个散射点进行复杂的运动补偿。

2 瞬时距离像成像

由于目标高速径向运动将导致解线调后的频谱产生搬移和扩展,由单频信号变为线性调频信号。传统思路是进行速度补偿,消除高速带来的影响,本文采用一种求瞬时距离像的方法,即对每次脉冲回波做时频分析,把各次脉冲回波的时频分布在同一采样时刻切片排列在一起,再经过运动补偿,横向瞬时多普勒分析,最终得到目标二维像。

传统的时频分析方法分为线性和双线性两种,线性时频分布主要有短时傅里叶变换、小波变换等;双线性变换有 Wigner-Ville 分布,模糊函数等。对每次脉冲进行时频分析时,双线性变换自项相乘会导致回波相位信息丢失,造成不能进行横向分辨,因此不适合做瞬时距离像分析。短时傅里叶变换是一种线性变换,它很好的保留了横向相位信息同时又不会受到交叉项的影响,对于空间轨道目标,径向高速运动带来一维距离像的搬移和扩展,但斜率均相同,且很小,各分量的最佳窗长应该是相同的,即局部平稳长度相同,因此,找到其中一个分量的最佳窗,则整个信号的最佳窗也就确定了,这里可以采用最小熵判别方法自适应地获得每个脉冲内各分量的最佳窗长,而且考虑到信号斜率很小,斜率缓慢变化,其最佳窗较长,在窗内信号近似平稳,这时可以达到很好的斜率分辨率,所以采用短时傅里叶变换和最小熵准则求最佳窗宽的方法来得到目标一维距离像。详细处理过程请参看文献[6]。

3 瞬时多普勒成像

3.1 从Foueiier变换到时频分析

Fourier 变换虽然可以将信号从时域变换到频域,从整体上将信号分解为不同的频率分量,但是对随时间频率变化的机动目标回波信号,其不能刻画某种频率分量发生在哪些时间内。而时频分析方法可以同时描述信号随时间变化的频谱变化。时频分析的基本思想是设计时间和频率的联合函数,用它同时描述在不同时间和频率的能量密度或强度。

例如,图 1(a)中,下面底图是 LFM 时域波形,右边图是频域谱图,左上方是时频分布图。容易看到,仅仅从谱图不能看出频率如何随时间变化,从时频分布图不仅可以看出频率随时间如何变化,而且从相对亮度可以看出频率强度。而机动目标的频率是时变的,如图 1(b)是

雷达实测数据某个距离单元散射点的时频分布，我们发现散射点的频率随时间无规律变化，如果利用基于 Fourier 变换的 RD 算法成像，会使散射点展宽图像模糊。由于时频分析可以得到目标的瞬时像，不会将目标同一散射点不同时刻的不同频率叠加在一起，所以利用时频分析可得到目标的清晰的像。

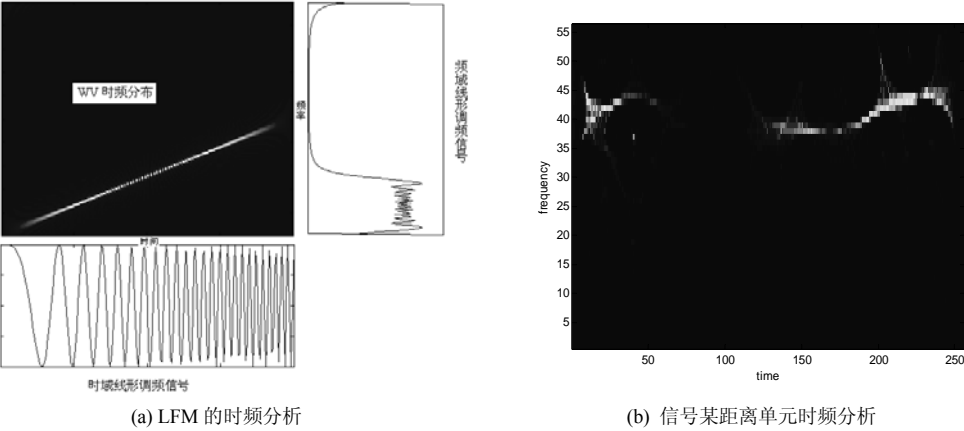


图 1 时频分析示例

通过相关的理论分析^{[7][9][10]}和我们的仿真分析，总的来说，时频集聚性最好的魏格纳-维尔分布(WVD)能获得高分辨的距离-瞬时多普勒像，但这种分布的主要缺点是存在交叉项干扰。目标的任一个距离单元内一般都有多个散射点，属多分量情况，对某一时刻切片的瞬时多普勒谱，交叉项作为干扰会污染图像。平滑伪魏格纳-维尔分布(SPWVD)是对 WVD 进行平滑处理对交叉项起了抑制作用，但它却以降低分布的时频分辨率为代价。短时傅立叶变换、Wigner-Ville、Choi-Williams、Margenau-Hill 谱图等时频分布存在时频聚集性差或交叉项干扰等影响，致使没有好的成像效果，而重排 Gabor 谱时频分析不仅有好的时频聚集性，而且没有交叉项干扰，所以我们选择重排 Gabor 谱时频分析对机动目标成像。

3.2 瞬时多普勒成像

首先再强调一下应用时频分布成像的原因。当目标机动飞行时，各个散射点的多普勒频率不是常数，而是随时间变化，此外目标散射点还会产生距离单元游动，目标图像就会变得模糊。用傅氏变换进行谱分析得到的是某一段时间内信号所包含的频率，由于目标机动飞行时各散射点回波多普勒谱是时变的，所以用时频变换代替傅氏变换，分析信号不同时刻的频率，从而得到目标的距离-瞬时多普勒像。用时频分析方法不仅可以得到清晰的目标图像，而且不需要进行复杂的运动补偿，不必采用能精确到任意精度的频域相关法和更复杂的如秩一相位估计(ROPE)等相位补偿法，文中包络对齐采用的是全局相关法，相位补偿采用的是多普勒中心跟踪法就可以得到满意的成像效果。在众多的时频分布中，重排 Gabor 谱时频分布不仅时频聚集性极好，而且完全不存在交叉项，这样成像分辨率不仅高，而且不受干扰，ISAR 成像质量很高，所以选择重排 Gabor 谱时频分布进行 ISAR 成像。

3.3 高速机动目标ISAR成像

利用瞬时距离-瞬时多普勒技术对高速机动目标 ISAR 成像流程如图 2。首先，对雷达回波信号利用短时傅里叶变换作为时频分析进行距离像压缩得到瞬时距离像，然后对其运动补偿（包括包络对齐和相位校正），运动补偿这部分大家都很熟悉，利用时频分析进行瞬时多普勒成像与常规成像处理的主要差别就是用时频变换代替 FFT。具体做法是对每一距离单元的数据进行重排 Gabor 谱时频变换，则会得到时间-多普勒二维数据，再联合距离维，形成时间-距离-多普勒三维图像方阵，沿着每一瞬时的距离-多普勒切片就是该时刻对应的 ISAR 图像，由于在各个瞬时散射点的多普勒值是固定的，因此不会造成图像的模糊。这种方法反映了 ISAR 图像随时间的变化规律，对于多普勒补偿要求极大地降低了，也能解决方位角依赖的发射性散射点在图像横向距离维造成的模糊。

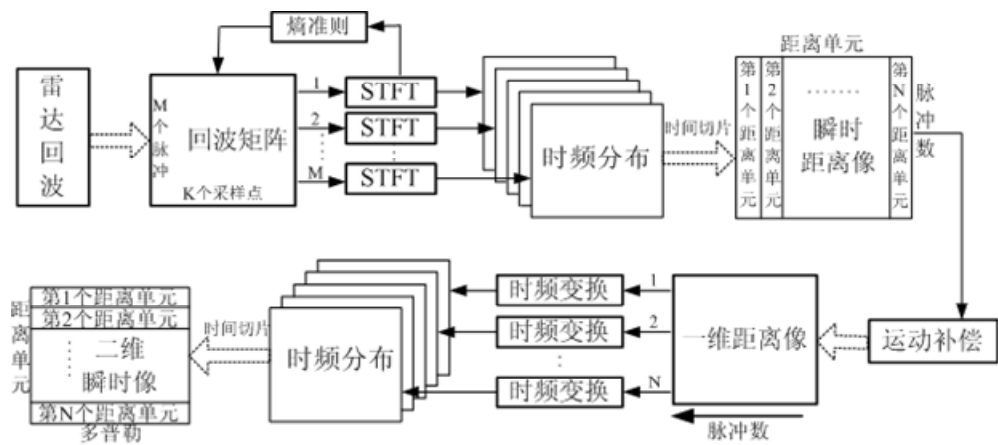


图 2 机动目标时频 ISAR 成像流程示意图

4 仿真试验及分析

为了验证成像算法的有效性，首先以高速机动飞行的导弹仿真数据验证，导弹模型如图 3 (a) 所示，由 16 个点组成，数据的仿真条件是：雷达中心频率 10G，带宽 1.7G，重频为 150Hz，其中导弹作高速且有机动的飞行：雷达位于原点，初始飞行速度 $v_0=[3500,0,0]$ ，观测起始时刻目标速度与雷达视线夹角为 45° ，考虑到可能的随机机动性，导弹飞行速度变为 $v_t=[3627.2,137.6,123.5]$ ，这时的径向速度也在变化，即存在机动性。脉内采样点数为 1024 个，横向积累脉冲数为 256 个，达到与距离分辨相匹配的转角，图 3 (b-d) 是利用导弹的仿真数据用各种算法成像结果。由于导弹存在高速和机动性，传统的 RD 成像算法使目标散射点展宽，图像模糊（图 3(b)）；距离瞬时成像方法虽然在距离向高速性得到补偿，但机动性使得时频聚集性差，即散射点展宽，时频分辨力降低（图 3(c)）；本文所提出的方法不仅能解决目标高速性带来的影响，而且目标的机动性带来的图像模糊也得到很好的处理，时频聚集性优良，而且无交叉项干扰，得到清晰的 ISAR 像（图 3(d)），这与理论分析是一致的。

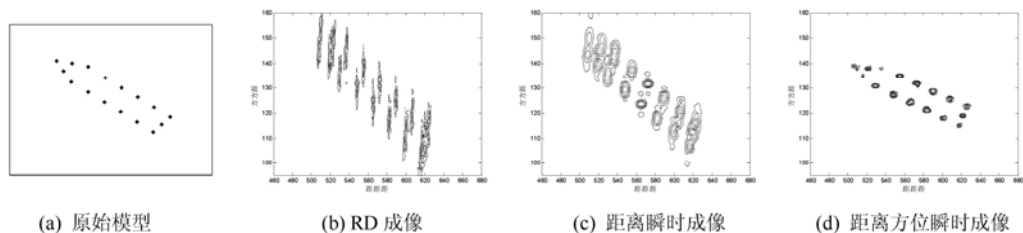


图 3 导弹模型及不同方法成像结果

同样，图 4 是利用飞机仿真数据用不同方法成像的结果。该数据的仿真条件是：雷达中心频率 9G，带宽 1G，重频为 150Hz，由 46 个点组成的飞机散射点模型如图 4 (a)，其中飞机作机动飞行描述如下：雷达位于原点，飞行中心点坐标 $R_{post}=[1500,1500,500]$ ，初始飞行速度 $v_0=[200,0,0]$ ，通过 1.67 秒的随机增速机动飞行，飞机飞行速度变为 $v_t=[325.5,124.5,124.0]$ 。容易发现，由于飞机或多或少存在机动性，目标方位角频率已不能近似为常数，传统 RD 成像方法会使目标轮廓展宽模糊，其他时频分析方法存在时频聚集性差或交叉项干扰的缺陷，采用重排 Gabor 时频分析瞬时成像可以得到目标清晰的像。验证了与上面同样的结论。

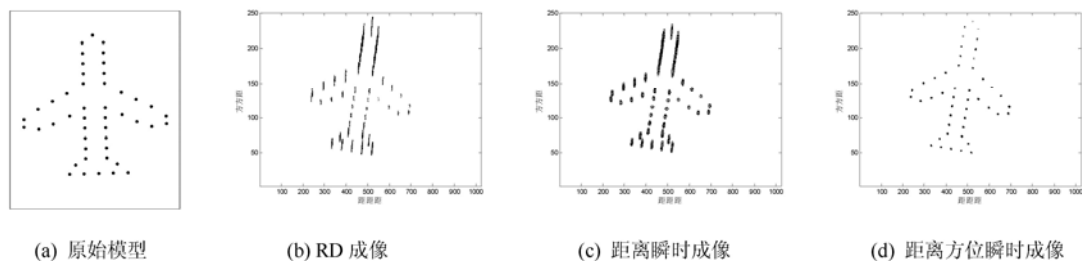


图 4 导弹模型及不同方法成像结果

5 结束语

针对傅里叶变换脉压对高速目标一维像的畸变和传统 RD 算法对机动目标成像的缺陷，本文采用短时傅里叶变换解决目标高速性，利用重排 Gabor 时频分布优良的时频聚集性和不受交叉项干扰特点，对目标瞬时成像，效果较好，最后通过仿真实验验证了所提出算法的正确性和潜力。

参 考 文 献

- [1] Martorella M, Acito N, Berizzi F. Statistical clean technique for ISAR imaging [J]. IEEE Trans. on Geoscience and Remote Sensing, 2007, 45(11): 3552-3560.
- [2] Hajduch G, Le Caillec J M, Garello, R. Airborne high-resolution ISAR imaging of ship targets at sea [J]. IEEE Trans. on Aerospace and Electronic Systems, 2004, 40(1): 378-384.
- [3] Musman S, Kerr D, Bachmann C. Automatic recognition of ISAR ship images [J]. IEEE Trans. on Aerospace and Electronic Systems, 1996, 32(4): 1392-1404.
- [4] 黄小红, 邱兆坤, 王伟. 目标高速运动对宽带一维距离像的影响及补偿方法研究 [J]. 信号处理, 2002, 18(6): 487 - 490.

- [5] Liu Aifang, Zhu Xiaohua, Lu Jinhui. The ISAR range profile compensation of fast-moving target using the dechirp method [A][C], IEEE Int. Conf. Neural Network and Signal Processing 2003, pp. 1619 – 1623.
- [6] Zhang Yi'an, Yin Zhiping, Chen Weidong, Wang Dongjin. Two-dimensional Radar Imaging Based on Instant Range Profile [A][C], 2007 1st Asian and Pacific Conference on Synthetic Aperture Radar Processing, pp. 236 – 239.
- [7] Shie Qian, Dapang Chen. Joint time-frequency analysis: method and applications [M]. London: Prentice-Hall International (UK) Limited, 1996.
- [8] Chen V C, Shie Qian. Joint time-frequency transform for radar range-Doppler imaging [J]. IEEE Trans. on Aerospace and Electronic Systems, 1998, 34(2): 486-499.
- [9] 张贤达, 保铮. 非平稳信号分析与处理 [M]. 北京: 国防工业出版社, 1998.
- [10] Kennedy S, Thayaparan T. Application of joint time-frequency representations to a maneuvering air target in sea-clutter: analysis beyond FFT [R], defence research and development Canada, ottawa, 2003, 49 Pages(s).

作者简介:

句彦伟 (1978—), 男, 山西朔州人, 博士, 南京电子技术研究所工作, 工程师, 研究方向为雷达信号处理

句彦伟 E-mail: juyanwei@126.com

通信地址: 南京市雨花区国睿路 8 号 南京 3918 信箱 60 分箱 邮编 210039

联系电话: 15950578192

基于Markov随机场和模拟退火算法的SAR 图像噪声抑制

钱 博

(中国电子科技集团南京电子技术研究所, 江苏南京, 2100013)

摘 要: 合成孔径雷达具有全天候监测能力, 受天气条件的约束更小, 成为重要的信息获取设备。但由于其成像原理导致的相干斑噪声问题大大影响了图像的质量, 因此, 去噪技术成为 SAR 图像处理的研究热点之一。本文提出了一种基于 Markov 随机场的噪声模型, 并以此建立目标函数。接着采用 Hamersley-Clifford 理论导出能量函数, 利用模拟退火算法优化目标函数, 实现噪声去除。最后, 本文利用均方误差、信噪比、边缘保持系数和等效视数四个参量将新算法与经典的增强 Lee 算法和 Frost 算法进行对比。实验结果表明, 本文提出的算法能更好地抑制噪声、保持图像细节。

关键词: SAR 图像处理; 斑噪; Markov 随机场; 模拟退火算法

SAR Image Denoising Based on Markov Random Field and Simulated Annealing

Qian Bo

(Nanjing Research Institute of Electronic Technology, Nanjing, Jiangsu, 210013, China)

Abstract: Synthetic aperture radar (SAR) remote-sensing instruments with the capabilities of all weather and all day/night, penetration, and terrain detection, provide unique geological environmental information, quite different from, yet complementary to passive systems such as visible remote sensing. However, the drawback of this autonomy is their higher susceptibility to speckle noise, so many researchers pay attention to speckle reduction. This paper proposed a method to reduce speckle noise in SAR image, Markov random field (MRF) was used to model the speckle and construct target function. Hamersley-Clifford theory was then used to derive the energy function associated with the MRF. After that, Simulated Annealing algorithm was employed for speckle reduction. Four metrics were used to assess the quality of the speckle reduction: the mean-square error, SNR, the edge-preservation parameter and the equivalent number of looks. The experiment result indicated that our proposed method performs better than enhanced Lee filter and Frost filter, and preserved more image details.

Keywords: SAR image processing; Speckle; Markov Random Field; Simulated Annealing Algorithm

1 引言

合成孔径雷达（SAR）可以实现全天候、全天时的监视工作，并且可以对各种目标进行高分辨率成像，在军事侦察监视以及地球物理探测、地面观测等多个领域都有着很大的应用潜力，受到广泛的关注。然而，SAR 图像中由其成像机制带来的相干斑噪声，无法从本质上消除^[1]；它对图像的后续处理，如目标辨识、目标检测、分类识别和各种应用会产生很大的影响。因此降噪问题成为 SAR 图像处理领域一个重要的研究课题（也有专家认为斑噪也是 SAR 图像信息的一部分，不需要进行去噪处理，因此目前存在抑制斑噪后处理图像和直接建立模型处理图像两种思路）。本文将沿袭去除斑噪后进行目标处理和检测的思路，重点考虑噪声抑制算法。

目前已有的相干斑噪声抑制算法主要分为空域滤波、变换域滤波和优化方法三类^[2,3]。其中，经典的空域滤波算法包括 Lee 滤波、Kuan 滤波、Frost 滤波、Gamma MAP 滤波及它们的增强算法。变换域滤波则将空域信号转变为频域信号后进行处理，主要的研究方向包括频谱分析和小波分解处理等。优化方法则根据已有的优化算法实现噪声的判决，如拓扑理论、模糊理论等。

近年来，马尔可夫随机场由于其强大的图像数据邻域空间相关性描述能力，以及优化算法的支持，成为 SAR 图像处理发展的新动力^[4]。

本文算法的思想是利用 Markov 随机场的空域描述能力建立斑噪的模型，建立目标函数；通过 Hamersley- Clifford 理论将目标函数转换为可以进行优化处理的能量函数，最后利用模拟退火算法的全局搜索能力得到能量函数的最优解，实现对噪声的抑制。实验结果表明，从均方误差、信噪比、边缘保持系数和等效视数四个参数来看，这种方法可以比增强 Lee 滤波和 Frost 滤波更有效地抑制斑噪的影响，保留目标的结构和纹理。

2 相关理论和系统模型

2.1 马尔可夫随机场

一幅 SAR 图像可以看作二维点集 $S = \{s = (i, j) | 1 \leq i \leq M, 1 \leq j \leq N\}$ ，其中 M 和 N 分别表示图像的宽度和高度。令 $F = \{F_i | i \in S\}$ 为定义于集合 S 的随机变量集合，每个随机变量 F_i 的一次取值为 $f_i, i \in S, f_i \in L$ ，其中 L 为随机变量 F_i 所有可能取值的集合，称为标识集， F 被称为随机场。若满足下列条件，则 F 被称为定义在 S 上邻域系统为 N 的马尔可夫随机场：

- 1) $P(f) > 0, \forall f \in F$;
- 2) $P(f_i | f_{S \setminus i}) = P(f_i | f_{N_i})$

其中 $S \setminus i$ 表示 S 中 i 的补集， $f_{N_i} = \{f_{i'} | i' \in N_i\}$ ，其中 N_i 是位置 i 的邻域集合。邻域集合满足： $i \notin N_i$ ，即当前位置不属于其邻域； $i \in N_{i'} \Leftrightarrow i' \in N_i$ ，即邻域关系是对应的。

由此可以得到一、二阶马尔可夫邻域模型如图 1 所示：

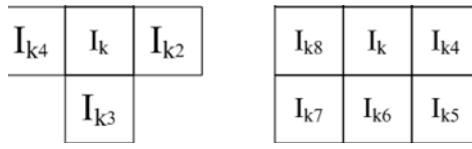


图 1 一、二阶马尔可夫邻域模型示意图

对于一阶马尔可夫模型，根据 Goodman 的统计最优理论，可以得到中心像素与邻域像素的关系如式 1、2 所示^[1]。

$$p_{I_k|I_{k1...4}}(i_k | i_{k1...4}) = \exp[-U(i_k | i_{k1...4})] \quad (1)$$

$$U(i_k | i_{k1...4}) = V_{C_1}(i_k) + V_{C_2}(i_k | i_{k1...4}) \quad (2)$$

根据 Hammers-Clifford 定理^[5]，可以得到 Markov 随机场和吉布斯分布之间具有同一性。根据这个性质，可以用吉布斯描述马尔可夫随机场。吉布斯分布可以用下式描述：

$$P(X=x) = \frac{1}{Z} e^{-U(x)/T} \quad (3)$$

其中 $Z = \sum_{x \in X} e^{-U(x)/T}$ 且 $U(x) = \sum_{c \in N} V_c(x)$ 。从此式可以看出，当温度 T 不变时，随机场分布的能量越小则分布的概率越大。当温度升高时， $P(X=x)$ 越平坦，可能出现的各种随机配置更多；反之能量配置出现的概率越大。以此作为能量目标函数，就可以利用模拟退火算法实现优化计算。

2.2 模拟退火算法

对于吉布斯能量函数，模拟退火算法能够实现全局搜索，得到能量函数的最小值。模拟退火算法的三个基本要素是邻域结构和新解生成法则、Metropolis 接受准则和冷却进度表。针对本文噪声抑制目标，为了实现目标能量函数的优化，即实现噪声的抑制和目标结构信息的保持，采用结合模拟退火和 Metropolis 采样的算法。优化算法详细步骤如下：

Step 1: 初始化，设置 Markov 随机场模型的初始值；

Step 2: 归一化测试，给定像素 k 的灰度为 i_k ：

测试是否为噪声点，若不是噪声点，保留原灰度不变；否则进入 Step3；

Step 3: 灰度值更新：

1. 根据邻域模型中其他像素的灰度产生新灰度值 i_{new} 。

2. 根据公式 $T_k = \lambda \times T_{k-1}$ 更新温度。

3. 计算 $p = \min\{1, e^{-\Delta U/T_k}\}$ ， $\Delta U = U(i_{new}, i_{k1...4}) - U(i_k, i_{k1...4})$ 。

4. 随机产生 $r \in [0,1]$ ；若 $r < p$ ，接受 i_{new} ；若 $r > p$ ，拒绝 i_{new} 。

Step 4: 以下一个像素为中心，重复 Step2 和 Step3。

Step 5: 当不是噪声点的像素占总像素数的比例大于给定阈值时，结束迭代。

Step2 到 Step4 实现了 SA-MS 算法的一次迭代。并通过非噪声点的覆盖范围控制图像迭代的结束。这种处理方式利用全局特征实现噪声点检测，并通过局部特征保留了纹理和结构特征。

2.3 SAR图像质量的评价

对于噪声处理技术来说，降噪前后的图像需要一些定量描述反映图像的质量，从而分辨处理算法的效果。本文将采用四类参数。

1) 均方误差 (MSE)

$$MSE = K^{-1} \sum_{i=1}^K (I_i - \hat{I}_i)^2$$

其中， K 表示图像中总的像素数， \hat{I}_i 表示降噪后像素 i 的灰度值， I_i 表示原图像中像素 i 的灰度值。

2) β 参量，主要用于反映边缘特征的保持程度

$$\beta = \frac{\Gamma(I_H - \bar{I}_H, \bar{I}_H - \hat{\bar{I}}_H)}{\sqrt{\Gamma(I_H - \bar{I}_H, I_H - \bar{I}_H) \Gamma(\bar{I}_H - \hat{\bar{I}}_H, \bar{I}_H - \hat{\bar{I}}_H)}}$$

其中 $\Gamma(I_1, I_2) = \sum_{i=1}^K I_{1_i} I_{2_i}$ ，若边缘完全保持，则 $\beta=1$ 。

3) 信噪比 (SNR)

$$SNR = 10 \log_{10} \left(\sum_{j=1}^K I_i^2 / \sum_{j=1}^K (I_i - \hat{I}_i)^2 \right)$$

4) 等效视数 (ENL)，SAR 图像处理中常用的斑噪水平描述参量。该值越高，对应区域内斑噪能量越小。

$$ENL = (mean^2)_{UniformArea} / (variance)_{UniformArea}$$

2.4 本文算法系统模型

综合上述介绍可以得到本文系统如图 2 所示：对于原始图像，通过马尔可夫随机场建模获得条件概率密度函数，建立噪声和信息模型。根据 Hammers-Clifford 定理将条件概率密度函数转化为吉布斯分布，形成能量函数。最后利用模拟退火算法优化能量函数，得到最优解即降噪的图像。

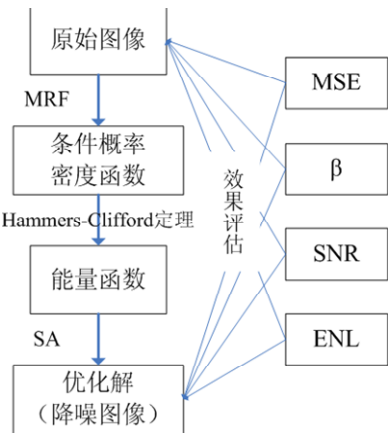


图 2 基于 MRF 和 SA 的噪声抑制算法示意图

实验过程中，本文将利用均方误差、边缘保持系数、信噪比和等效视数四个参数对降噪后得到的图像质量，以及降噪算法性能进行评估。并与经典的斑噪抑制算法--增强 Lee 和 Frost 算法进行对比。

3 实验、结果与分析

本文主要利用实验图像为利用光学图像为输入构造的乘性噪声仿真图像和实际机载 SAR 图像进行两项实验。第一项实验按信噪比将乘性噪声叠加到光学图像上模拟实际成像中产生的斑噪，再利用几种去噪算法（本文算法、增强 Lee 滤波和 Frost 滤波）得到去噪后的图象，观察滤波效果、对比评估参数。第二项实验则应用实际 SAR 图像观察算法性能。

3.1 构造Speckle噪声图像及去噪实验

图 3 为仿真噪声图像及降噪效果实验图，其中(a)为原始光学图像，(b)为叠加噪声的图像，(c)为叠加的噪声(基准噪声图像)，(d)为 5*5 增强 Lee 滤波算法滤除的噪声图像，(e)为 5*5Frost 滤波算法滤除的噪声图像，(f)本文算法 50 次迭代滤除的噪声图像。

从滤除的噪声图像可以更为清晰地看出算法对噪声的抑制能力：由于添加的 Speckle 噪声是乘性噪声，噪声的能量与原图的灰度结构有很大的关系，因此滤除的噪声图像应该包含一定的类似原图灰度结构的纹理，但是如果算法过于平滑，会把原始图像中目标包含的边缘信息滤除掉，使图像变模糊。

从图中可以看出，增强 Lee 滤波算法可以相对平滑地从含噪图像中消除噪声，携带了部分边缘信息；Frost 滤波算法的处理得结果去掉了原始图像中边缘保持的能量，这样使原图的结构信息受到损失。而本文算法得到的噪声图像更接近于叠加进去的噪声，这表明本文算法不仅对 Speckle 噪声有较强的噪声抑制能力，而且对原图边缘信息的保持也做得比较好。因此，从仿真图像上分析，本文算法效果最好。

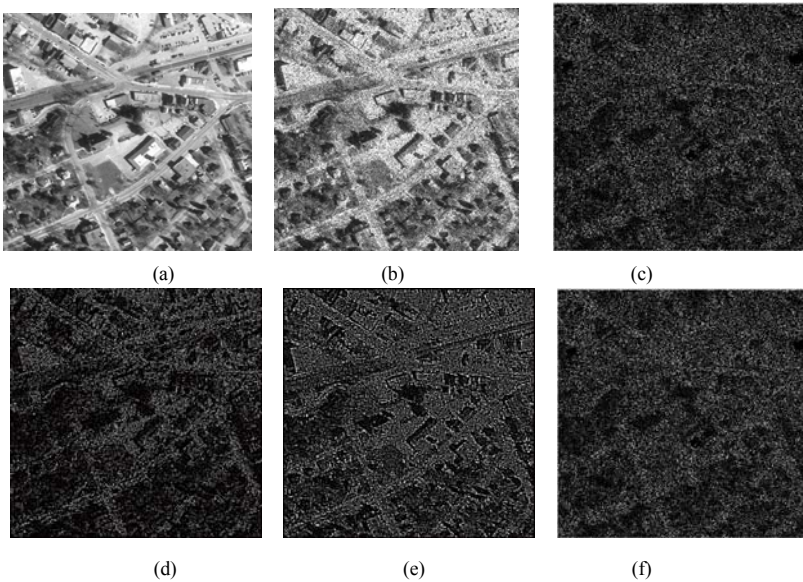


图 3 滤波算法滤除地噪声图像对比

此外，表 1 中给出了几种去噪算法的评估参数，可以从定量的角度对比几种算法的性能。从表中数据可以看出，本文算法在等效视数、均方误差、边缘保持能力和信噪比四个评估指标上都表现了较强的性能，超过了经典地增强 Lee 滤波和 Frost 滤波降噪算法。不仅如此，处理效果和退火算法的迭代次数也有着密切的关系。实验表明，通过增加迭代次数可以提高处理的效果，但系统的性能会渐渐趋向于稳定值。考虑到迭代次数的增加会加大计算密度，在实际应用中如何确定合适的迭代次数需要通过进一步的实验和理论分析确定。

表 1 系统推荐成功率（%）对比和分析

	ENL	MSE	β	SNR(dB)
含噪图像	13.6	967.2	0.21	15
增强 Lee	31.2	387.5	0.36	19.2
Frost	35.4	382.6	0.31	19.4
20 次迭代	49.1	246.9	0.44	23.6
50 次迭代	60.7	221.1	0.46	26.8

3.2 实际SAR图像去噪实验

本实验将对实际的 SAR 图像进行本文算法的处理，采用的迭代次数为 50 次，原图像和滤波后的实际效果如图 4 所示。

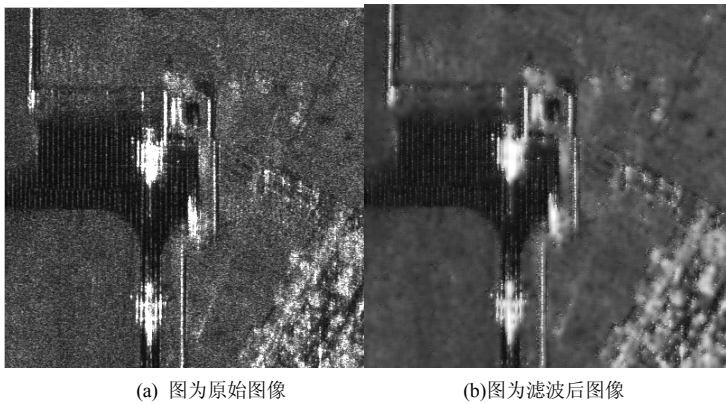


图 4 本文滤波算法效果图

从图 4(a)(b)中可以看出，对于大面积灰度较稳定的区域（一般对于农田、平地，如左下角草坪区域），噪声被平滑处理，只保留了平均能量信息。对于建筑物群（如右下角居民区）则进行了灰度修正，保留了纹理信息。对于具有大量灰度阶越的目标区域则尽量保持原有的边缘信息。总体看来，经过降噪处理后，虽然从人眼观察角度感觉图像有些模糊，但是从计算机处理的角度来看，目标和背景可以更好地通过纹理特征、边缘特征和灰度特征进行分割，这样更有利于后续目标分类和识别算法的设计和实现。

4 结论和展望

本文针对 SAR 图像噪声处理问题提出了一种新的算法，该算法利用马尔可夫随机场对

SAR 图像中的噪声进行建模, 得到概率密度函数后, 根据 Hemmers-Clifford 理论转换为能量目标函数, 最后用模拟退火算法进行优化实现噪声抑制。从仿真实验数据来看, 各项指标上本文算法都优于增强 Lee 滤波和 Frost 滤波法, 对实际 SAR 图像也能够实现抑制噪声保护目标纹理和结构信息的要求。此外, 随着模拟退火算法迭代次数的增加, 算法还会进一步提高降噪的性能, 但这同时也会增加计算开销。如何在性能和计算负担之间寻找平衡成为新的问题, 后续的工作将着眼于从图像信息分布特征上寻找加快退火算法处理速度等方面。

参 考 文 献

- [1] Goodman J W. Some fundamental properties of speckle [J]. Journal of the Optical Society of America, 1976, 66(11): 1145-1150.
- [2] Lee J S. Digital image enhancement and noise filtering by use of local statistics[J]. IEEE Trans on Pattern Analysis and Machine Intelligence. 1980, 2(2): 165-168.
- [3] Lopes A, Touzi R, Nezry E. Adaptive speckle filters and scene heterogeneity[J]. IEEE Trans on Geoscience and Remote Sensing, 1990, 28(6): 992-1000.
- [4] C. Oliver and S. Quegan, Understanding Synthetic Aperture Radar Images, SciTech, NC 27613, 2004.
- [5] J. Besag, Spatial interaction and the statistical analysis of lattice system, J. R. Stat. Soc. Series B(36), pp. 192-236, 1974.
- [6] 候一民, 郭雷, 一种基于马尔可夫随机场的 SAR 图像分割新方法, 电子与信息学报, Vol. 29(5), pp: 1069-1072.
- [7] Lankoande O, Hayat M M, and Santhanam B. Speckle reduction of SAR images using a physically based Markov random field model and simulated annealing. Proc. of SPIE, 2005, Vol. 5808: 210-221.
- [8] 周荫清, 张磊, 徐华平, 小波模拟退火-软阈值 SAR 图像降噪算法, 北京航空航天大学学报, Vol. 32(3), pp: 293-296.

作者简介:

钱博, 男, 博士, 1981 年出生于江苏南京, 研究方向为 SAR 图像处理和目标识别。

循环相关谱检测方法及其门限的确定

赵知劲^{1,2} 张伟卫¹ 徐世宇¹

(1.杭州电子科技大学通信工程学院, 浙江 杭州 310018;

2.中国电子集团第36研究所通信系统信息控制技术国家级重点实验室, 浙江嘉兴, 314001)

摘要: 频谱感知是认知无线电的关键技术之一。针对循环平稳特性检测法计算复杂度高和检测门限难以确定, 本文提出只对循环相关谱特定的峰值点进行检测的循环相关谱检测法和循环相关谱系数检测法, 推导得到了这两种方法检测统计量的概率密度函数和检测门限。仿真结果表明, 循环相关谱检测法的检测性能优于循环相关谱系数检测法, 但循环相关谱检测法的性能受噪声功率估计误差影响大, 而循环相关谱系数检测法不需要已知噪声功率, 实用性更强。

关键词: 认知无线电; 频谱感知; 循环平稳特性; 检测门限

Detection Methods and Their Thresholds Based on Cycle Coherence Spectrum

Zhao Zhi jin^{1,2} Zhang Wei wei¹ Xu Shi yu¹

(1 Telecommunication School, Hangzhou Dianzi University, Hangzhou 310018, China; 2 State key lab of information control technology in communication system of No.36 Research Institute, China Electronic Technology Corporation, Jiaxing 314001, China)

Abstract: Spectrum sensing plays a key role in cognitive radio. Because the computation complexity of traditional cyclostationarity feature detector is high and the thresholds are hard to set, this paper proposes two methods which only detect on some peak points of cycle coherence spectrum, named a detection method based on cycle coherent spectrum and a detection method based on cycle coherent spectrum coefficient. And then the probability density functions of their detection statistics and thresholds are deduced. Simulation results show that the performance of the method based on cycle coherent spectrum is better than that of the method based on cycle coherent spectrum coefficient, but it is affected by the noise power greatly. The later doesn't need the prior information of noise power and is suitable for practical application.

Keywords: cognitive radio; spectrum sensing; cyclostationarity feature; detection thresholds

1 引言

无线频谱是一种宝贵的自然资源，但随着无线电通信及其应用的发展，频谱稀缺问题越来越突出。然而，根据美国联邦通信委员会的报告估计^[1]，目前 3GHz 以下授权频段的频谱利用率只有 15%到 85%，3~6GHz 频段的频谱利用率甚至不到 0.5%。显然，频谱资源并不是真正地短缺，而是存在大量频谱使用不均的问题。认知无线电^{[2][3]}通过感知外部一定频带范围内的频谱环境，利用暂时未被占用的频带进行机会式无线传输，从而能在不影响主用户正常使用的前提下动态地接入未被使用的频谱。频谱感知作为实现上述功能的关键技术之一而显得格外重要。

循环平稳特性检测法利用已调信号内在的周期性来检测信号^[4]。针对传统的循环平稳特性检测法计算复杂度高和检测门限难以确定的问题，为降低计算复杂度，本文只对循环相关谱的峰值点进行检测，提出循环相关谱检测法和循环相关谱系数检测法；推导了只有噪声存在时统计量的近似分布，给出了一定虚警概率下的检测门限。

2 循环谱检测法

由于通信信号通常要进行载波调制、过采样、冗余编码等操作，从而会产生内在的周期平稳性，当信号的均值和自相关函数均存在周期性，则称该信号是二阶循环平稳过程^[5]，其循环相关谱函数为

$$S_x^\alpha(f) = \int_{-\infty}^{+\infty} R_x^\alpha(\tau) e^{-j2\pi f\tau} d\tau = \frac{1}{T} X(f + \alpha/2) X^*(f - \alpha/2) \quad (1)$$

其中， $X(f)$ 为信号 $x(t)$ 的频谱。由上式可见，循环相关谱就是信号频谱的相关，噪声频谱不具有相关性，所以利用这个特性可以对低信噪比下的主用户信号进行检测。

当接收信号为经过采样的离散信号 $x(n)$ 时，对信号进行 N 点 FFT 并进行 M 次平滑可以得到 $S_x^\alpha(f)$ 的离散表达式为：

$$S_x^{2D}(k) = \frac{1}{MN} \sum_{m=1}^M X_m(k+D) X_m^*(k-D) \quad (2)$$

其中 $X_m(k)$ 为第 m 组数据 $x_m(n_1) = x((m-1)N/P + n_1)$ 的 N 点 FFT 变换， P 为数据的重叠因子， $k = f/\Delta F$ ， $D = \alpha/(2\Delta F)$ ，这里 $\Delta F = f_s/N$ 为两个相邻频率点之间的频谱间隔。

传统的循环平稳特性检测法计算复杂度高，难以满足认知无线电中的快速感知要求。本文假设主用户信号的调制方式和码元速率等参数已知，利用这些参数预先确定出主用户信号循环相关谱峰值点出现的位置，从而只需对循环相关谱在特定的峰值点上进行检测以降低复杂度，并称该方法为循环相关谱检测法。

传统的循环平稳特性检测法将 $|S_x^{2D}(k)|$ 作为检测统计量，由于循环相关谱就是信号频谱的相关，可以定义离散信号的循环相关谱系数为：

$$C_x^{2D}(k) = \frac{S_x^{2D}(k)}{[S_x(k+D) S_x(k-D)]^{1/2}} \quad (3)$$

其中, $S_x(k)$ 为信号 $x(n)$ 的功率谱。将 $|C_x^{2D}(k)|$ 作为统计量只在特定点上对主用户信号进行检测, 这就是本文提出的循环相关谱系数检测法。

3 检测门限的确定

3.1 循环相关谱检测法的检测门限

当主用户不存在时, 由式 (2) 可知 $S_x^{2D}(k) = S_w^{2D}(k) = \frac{1}{MN} \sum_{m=1}^M W_m(k+D)W_m^*(k-D)$, 其

中 $W(k)$ 为高斯白噪声 $w(n)$ 的离散傅立叶变换, $S_w^{2D}(k)$ 表示只有噪声存在时的循环相关谱。由于本文的目的是利用信号的循环平稳特性来检测主用户信号的存在与否, 所以这里只讨论 $D \neq 0$ 情形。当 $D \neq 0$ 时可以求得 $S_w^{2D}(k)$ 的期望为

$$E[S_w^{2D}(k)] = \frac{1}{MN} \sum_{m=1}^M E[W_m(k+D)W_m^*(k-D)] = 0 \quad (4)$$

令 $Z_m^D(k) = W_m(k+D)W_m^*(k-D)$ 。当噪声 $w(n)$ 为实高斯白噪声且 $k=0$ 时, 可以得到

$$\begin{aligned} E[Z_m^D(0)(Z_n^D(0))^*] &= E[W_m(D)W_m^*(-D)W_n^*(D)W_n(-D)] \\ &= E[W_m(D)W_m^*(-D)]E[W_n^*(D)W_n(-D)] + E[W_m(D)W_n^*(D)]E[W_m^*(-D)W_n(-D)] \\ &\quad + E[W_m(D)W_n(-D)]E[W_m^*(-D)W_n^*(D)] \\ &= \begin{cases} 2\sigma^4 N^2 (1 - \frac{|m-n|}{P})^2 e^{j4\pi D(m-n)/P} & |m-n| \leq P-1 \\ 0 & |m-n| \leq P \end{cases} \end{aligned} \quad (5)$$

其中 σ^2 为高斯白噪声的方差。由式 (2) 和式 (5) 可以得到 $S_w^{2D}(0)$ 的方差 V_{s1} 为

$$\begin{aligned} V_{s1} &= \text{Var}(S_w^{2D}(0)) = \frac{1}{N^2 M^2} E[(\sum_{m=1}^M W_m(D)W_m^*(-D))(\sum_{n=1}^M W_n(D)W_n^*(-D))^*] \\ &= \frac{1}{N^2 M^2} \sum_{m=1}^M \sum_{n=1}^M E[Z_m^D(0)(Z_n^D(0))^*] = \frac{2\sigma^4}{M^2} [M + \sum_{i=1}^{P-1} 2(M-i)(1 - \frac{i}{P})^2 \cos(4i\pi D/P)] \end{aligned} \quad (6)$$

可以证明, $S_w^{2D}(0)$ 的实部和虚部不相关且服从相同分布, 其均值为 0, 方差 σ_s^2 为 $V_{s1}/2$ 。由中心极限定理可得, $S_w^{2D}(0)$ 的实部和虚部分别近似服从 $N(0, V_{s1}/2)$ 分布。

当噪声 $w(n)$ 为实高斯白噪声而 $k \neq 0$ 时, 通过推导可以得到 $S_w^{2D}(0)$ 的方差 V_{s2} 为

$$V_{s2} = \text{Var}(S_w^{2D}(k)) = \frac{\sigma^4}{M^2} [M + \sum_{i=1}^{P-1} 2(M-i)(1 - \frac{i}{P})^2 \cos(4i\pi D/P)] \quad (7)$$

类似地, 可以得到 $S_w^{2D}(k)$ 的实部和虚部不相关且均近似服从 $N(0, V_{s2}/2)$ 分布。

当噪声 $w(n)$ 为复高斯白噪声时, 通过推导可以得出: $S_w^{2D}(k)$ 方差与 k 无关^[7], $S_w^{2D}(k)$ 的实部和虚部不相关, 且其分布与实高斯白噪声在 $k \neq 0$ 下的分布相同。

由上可知, 当 $D \neq 0$ 时, $S_w^{2D}(k)$ 为复数, 且实部和虚部均近似服从正态分布。由文献[8]可知, $|S_w^{2D}(k)|$ 的概率密度函数 (PDF) 为

$$p_s(r) = \frac{r}{\sigma_s^2} e^{-r^2/(2\sigma_s^2)} \quad (r \geq 0) \quad (8)$$

这里 $r = |S_w^{2D}(k)|$, σ_s^2 为 $S_w^{2D}(k)$ 的实部和虚部的方差。

当噪声 $w(n)$ 为实高斯白噪声且 $k = 0$ 时, $\sigma_s^2 = V_{s1} / 2$; 当噪声 $w(n)$ 为实高斯白噪声而 $k \neq 0$ 或复高斯白噪声时, $\sigma_s^2 = V_{s2} / 2$ 。当虚警概率 P_f 一定时, 利用式 (8) 可得循环相关谱检测法的门限值 Th_s 为

$$\text{Th}_s = \sqrt{2\sigma_s^2 \ln\left(\frac{1}{P_f}\right)} \quad (9)$$

3.2 循环相关谱系数检测法的检测门限

由式 (3) 可知, 只有噪声存在时的循环相关谱系数 $C_x^{2D}(k)$ 为

$$C_x^{2D}(k) = C_w^{2D}(k) = \frac{S_w^{2D}(k)}{[S_w(k+D)S_w(k-D)]^{1/2}} \quad (10)$$

其中 $S_w(k)$ 为高斯白噪声的功率谱密度, 其期望值为 σ^2 , $C_w^{2D}(k)$ 表示只有噪声存在时的循环相关谱, 可以得到

$$C_w^{2D}(k) = \frac{S_w^{2D}(k)}{\sigma^2} \quad (11)$$

由式 (6) 和式 (11) 可以得出当噪声为实高斯白噪声且 $k = 0$ 时, $C_w^{2D}(0)$ 的实部和虚部不相关, 且均近似服从 $N(0, V_{c1} / 2)$ 分布, 其中

$$V_{c1} = V_{s1} / \sigma^4 = \frac{2}{M^2} \left[M + \sum_{i=1}^{P-1} 2(M-i) \left(1 - \frac{i}{P}\right)^2 \cos(4i\pi D / P) \right] \quad (12)$$

当噪声为实高斯白噪声而 $k \neq 0$ 或复高斯白噪声时, 由式 (7) 和式 (11) 可得 $C_w^{2D}(k)$ 的方差 V_{c2} 为

$$V_{c2} = V_{s2} / \sigma^4 = \frac{1}{M^2} \left[M + \sum_{i=1}^{P-1} 2(M-i) \left(1 - \frac{i}{P}\right)^2 \cos(4i\pi D / P) \right] \quad (13)$$

类似地, 可以得到 $C_w^{2D}(k)$ 的实部和虚部不相关且均服从 $N(0, V_{c2} / 2)$ 分布。

由文献[8]可以得出 $|C_w^{2D}(k)|$ 的 PDF 为

$$p_c(r) = \frac{r}{\sigma_c^2} e^{-r^2 / (2\sigma_c^2)} \quad (r \geq 0) \quad (14)$$

其中 $r = |C_w^{2D}(k)|$, σ_c^2 为 $C_w^{2D}(k)$ 的实部和虚部的方差, 当噪声为实高斯白噪声且 $k = 0$ 时, $\sigma_c^2 = V_{c1} / 2$; 当噪声为实高斯白噪声而 $k \neq 0$ 或复高斯白噪声时, $\sigma_c^2 = V_{c2} / 2$ 。

当虚警概率 P_f 一定时, 由式 (14) 可得循环相关谱系数检测法的门限值 Th_c 为

$$\text{Th}_c = \sqrt{2\sigma_c^2 \ln\left(\frac{1}{P_f}\right)} \quad (15)$$

由式 (12) 和式 (13) 可见, V_{c1} 和 V_{c2} 的值与噪声功率 σ^2 无关; 再由式 (15) 可以得出, 该方法在一定虚警概率下的检测门限与噪声功率无关。

因此, 当 N 、 M 、 P 、 D 和 σ^2 已知时, 可以由式 (6)、式 (7) 和式 (9) 得到循环相关谱检测法在给定虚警概率下的检测门限 Th_s ; 同理, 当 N 、 M 、 P 、 D 已知时, 可以由式 (12)、式 (13) 和式 (15) 得到循环相关谱系数检测法在给定虚警概率下的检测门限 Th_c 。所以当噪声功率 σ^2 已知时, 两种检测法在给定虚警概率下的检测门限 Th_s 和 Th_c 都可以确定; 而当噪

声功率 σ^2 未知时, 检测门限 Th_s 无法确定, 但 Th_c 可以确定, 因此循环相关谱系数检测法更具有实用价值。

4 性能仿真与分析

对 $|S_w^{2D}(k)|$ 和 $|C_w^{2D}(k)|$ 分别进行 100000 次 Monte Carlo 仿真。取 $P = 1$ 、 $k = 0$, 图 1 和图 2 分别给出了仿真得到的 $|S_w^{2D}(k)|$ 和 $|C_w^{2D}(k)|$ 的概率密度函数曲线和由式 (8) 和式 (14) 计算得到的理论值曲线。由图 1 和图 2 可见, 仿真结果与理论分布很接近, 从而证明了式 (8) 和式 (14) 的正确性。

为了验证这两种方法的有效性, 对它们的性能分别进行仿真分析。本文选取码元速率 $R = 1/T_0 = 128$ 千符号/秒, 载波速率为 $f_c = 384\text{kHz}$ 的 BPSK 信号进行仿真, 其中采样频率 $f_s = 2048\text{kHz}$, FFT 的点数 N 取 1024。

(1) 循环相关谱检测法的检测门限

取虚警概率 $P_f = 0.1$, σ^2 分别取 1 和 1.4, 重叠因子 P 和组数 M 分别取 (1, 20) 时, 由式 (6)、(7) 和式 (9) 可得到 BPSK 信号利用循环相关谱检测法时在 $D = 192, k = 0$ 、 $D = 160, k = 0$ 、 $D = 32, k = 192$ (即 $\alpha = 2f_c f = 0$ 、 $\alpha = 2f_c - 1/T_0 f = 0$ 、 $\alpha = 1/T_0 f = f_c$) 这 3 个点上的检测门限值。表 1 分别给出了 σ^2 取 1 和 1.4 时的检测门限。

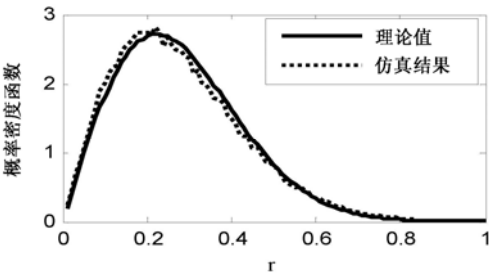


图 1 $|S_x^{2D}(k)|$ 仿真分布与理论分布的比较

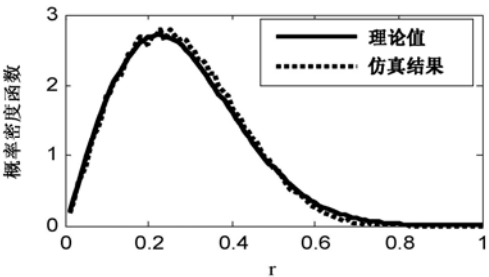


图 2 $|C_x^{2D}(k)|$ 仿真分布与理论分布的比较

表 1 $(P, M) = (1, 20)$ 时, 循环相关谱检测法的检测门限

σ^2	(D, k)		
	$(192, 0)$	$(160, 0)$	$(32, 192)$
1	0.48	0.48	0.34
1.4	0.67	0.67	0.476

(2) 循环相关谱检测法的性能

取 $P_f = 0.1$ 、 $M = 20$ 、 $\sigma^2 = 1$, P 分别取 1 和 8 时, 循环相关谱检测法对 BPSK 信号在 $D = 192, k = 0$ 、 $D = 160, k = 0$ 、 $D = 32, k = 192$ 这 3 个点进行检测的性能分别如图 3 和图 4 所示。由图 3 和图 4 可见, 当采用循环相关谱检测法时, 随着 P 的增加, 检测性能有所下降; 且 $D = 32, k = 192$ 点上的检测性能与 $D = 192, k = 0$ 点上的检测性能比较接近, 而 $D = 160, k = 0$ 点上的检测性能明显比其它两点差。

由于该方法需要根据噪声功率来确定门限, 当估计得到的噪声功率存在误差时, 将得到错误的门限值, 下面对这种情形进行分析。图 5 给出了 $P_f = 0.1$ 、 $M = 20$ 、 $P = 1$, 将噪声的方

差 $\sigma^2 = 1$ 错误的估计为 1.4 时利用循环相关谱检测法对 BPSK 信号在 $D = 192, k = 0$ 、 $D = 160, k = 0$ 、 $D = 32, k = 192$ 这 3 个点的检测性能。比较图 3 和图 5 可以发现，图 5 中 3 个检测点的性能明显下降，这是由于将噪声功率估计偏大而导致门限值偏大的缘故，由式 (9) 可得此时真实的虚警概率 P_f 为 0.04，小于 0.1。

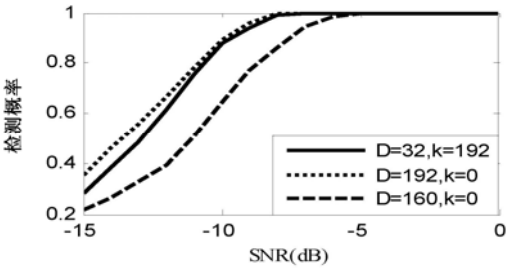


图 3 $P = 1$ 时基于循环相关谱检测法在各个检测点上的性能

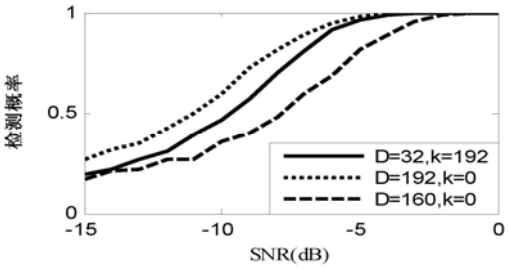


图 4 $P = 8$ 时基于循环相关谱检测法在各个检测点上的性能

(3) 循环相关谱系数检测法的检测门限

当虚警概率 P_f 为 0.1，重叠因子 P 和组数 M 分别取 (1, 20) 时，由式 (12)、(13) 和式 (15) 可得到 BPSK 信号的检测统计量 $|C_x^{2D}(k)|$ 在 $D = 192, k = 0$ 、 $D = 160, k = 0$ 、 $D = 32, k = 192$ 这 3 个点上的门限值与表 1 $\sigma^2 = 1$ 时相同。

(4) 循环相关谱系数检测法的性能

取 $P_f = 0.1$ 、 $M = 20$ 、 $P = 1$ ，在 σ^2 分别取 1 和 1.4 时利用循环相关谱系数检测法对 BPSK 信号在 $D = 192, k = 0$ 、 $D = 160, k = 0$ 、 $D = 32, k = 192$ 上的检测性能分别如图 6 和图 7 所示。由图 6 和图 7 可见，该方法在点 $D = 32, k = 192$ 上的检测性能略优于点 $D = 192, k = 0$ 上的性能，而在点 $D = 160, k = 0$ 上的检测性能相对较差，且 σ^2 取不同值时在同一门限下性能差别不大，从而表明该方法能克服噪声功率不确定性带来的影响。

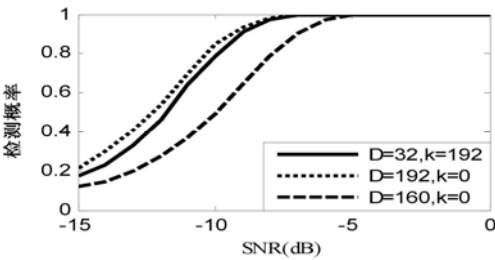


图 5 噪声功率估计偏大时循环相关谱检测法在各个检测点上的性能

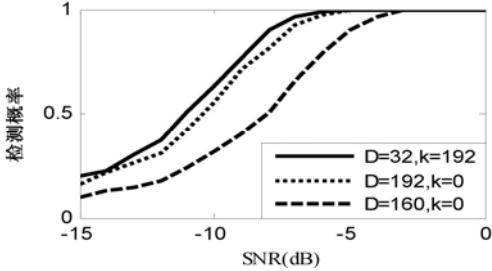


图 6 $\sigma^2 = 1$ 时，循环相关谱系数检测法在各个检测点上的性能

(5) 两种方法检测性能比较

图 8 给出了 $P_f = 0.1$ 、 $M = 20$ 、 $P = 1$ 时这两种方法在点 $D = 192, k = 0$ 、 $D = 32, k = 192$ 上的性能比较。由图 8 可见，循环相关谱检测法的性能优于循环相关谱系数检测法约 2~3dB。循环相关谱检测法需要已知噪声功率，而循环相关谱系数检测法不需要已知噪声功率，所以当噪声功率已知时，用循环相关谱检测法可以获更好的检测性能；而当噪声功率未知时，可以用循环相关谱系数检测法，该方法无需噪声功率大小作为先验信息。

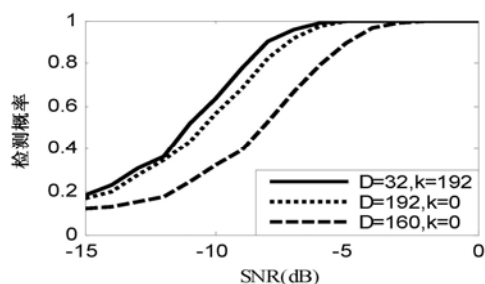


图 7 $\sigma^2 = 1.4$ 时, 循环相关谱系数
检测法在各个检测点上的性能

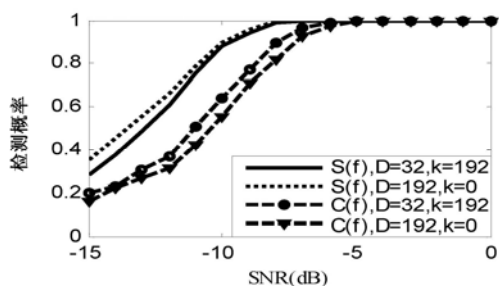


图 8 两种方法的检测性能比较

参 考 文 献

- [1] Federal Communications Commission. Spectrum Policy Task Force Report. Washington: ET Docket, 2002:2-135.
- [2] 赵知劲, 郑仕链, 尚俊娜. 认知无线电技术[M]. 北京: 科学出版社, 2008.
- [3] Akyildiz I F, Lee Q, Vuran M C. Next Generation/dynamic Spectrum Access/cognitive Radio Wireless Network: A Survey [J]. Computer Networks, 50(13), pp: 2127-2159, Sep 2006.
- [4] Xu S Y, Zhao Z J, Shang J N. Spectrum Sensing Based on Cyclostationarity [J]. Power Electronics and Intelligent Transportation System, pp: 171-174, Aug 2008.
- [5] 陈洸. 基于循环谱相关的调制方式识别的研究[D]. 北京交通大学硕士论文, 2006.
- [6] Kim K, Akbar I A, Bae K K. Cyclostationary Approaches to Signal Detection and Classification in Cognitive Radio[C]. Proc IEEE New Frontiers in Dynamic Spectrum Access Networks, DySPAN'07:212-215, Apr 2007.
- [7] Chen H S, Gao W, Dayt D G. Spectrum Sensing Using Cyclostationary Properties and Application to IEEE 802.22 WRAN[C]. Global Telecommunications Conference, GLOBECOM'07:3133-3138, Nov 2007.
- [8] 张力军, 张宗橙, 郑宝玉. 数字通信[M]. 北京: 电子工业出版社, 2001.

作者简介:

赵知劲(1959-), 女, 浙江宁波人, 博士, 教授、博士生导师, 通信信号处理、自适应信号处理、认知无线电技术等;

张伟卫(1986-), 男, 浙江台州人, 在读研究生, 认知无线电, Email: wz21103@gmail.com;

徐世宇(1984-), 男, 浙江台州人, 硕士, 认知无线电。

GPS信号捕获中信噪比提高的研究

雷 蕾 林静然 彭启琮

(电子科技大学 通信与信息工程学院, 成都 611731)

摘 要: 对 GPS 弱信号进行捕获时, 通常采用相干积分和非相干积分来提高信噪比。详细分析了相干积分、非相干积分、相干/非相干积分组合的增益及性能, 介绍了根据增益要求计算不同相干/非相干积分组合方式的方法, 并用真实采集的 GPS 信号对不同组合方式进行了仿真验证。在粗略知道 GPS 信号强度的情况下, 可以根据需要的信噪比计算出合适的组合方式, 来实现弱信号的捕获。

关键词: GPS; 捕获; 信噪比; 相干/非相干积分组合

Research on Improving SNR in Acquisition of GPS Signals

Lei Lei Lin Jing ran Peng Qi cong

(School of Communication and Information Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China)

Abstract: Coherent integration and non-coherent integration are always used to improve SNR in the acquisition of weak GPS signals. The gain and performance of coherent integration, non-coherent integration, and the combination of them are analyzed in this paper. The method of calculating different combinations of them based on the request gain is introduced, and the effects of these combinations are proved in the simulation tests based on the real GPS signals. In the acquisition when only the approximate intensity of GPS signals is available, the method can be used to find the proper combination according to the SNR needed.

Keywords: GPS (Global Positioning System); acquisition; SNR; combination of coherent integration and non-coherent integration

1 引言

随着定位服务需求的增加, GPS 开始被应用到一些复杂信号环境下, 如街道、室内及高速交通工具等。在这些环境下, GPS 信号受到严重的衰减, GPS 信号功率(一般低于-170dBW)远小于 GPS 接收机的正常工作范围(-160dBW 或-130 dBm)^[1]。普通的 GPS 接收机无法工作在这样环境, 必须采取措施来提高信噪比。

提高信噪比通常有两种方法: 相干积分和非相干积分(coherent integration and non-coherent integration)^[2,3], 有时也将两种积分方法结合来获得更高的信噪比^[4]。本文首先介绍了相关积分和非相干积分的增益和性能, 然后对两种积分累加结合对信噪比的影响进行了详细的分析。

最后在此基础上介绍了一种方法，在粗略知道信号强度的情况下，确定合适的相干/非相干积分组合方式来进行捕获。

2 相干/非相干积分组合方式的确定

2.1 相干积分和非相干积分简介

相干积分是将不同周期的接收 C/A 码信号与本地码信号的相关值进行对应叠加。非相干积分是对每个周期的接收 C/A 码信号与本地码信号的相关值取模，再将多周期的取模结果进行对应叠加^[5]。分别如公式(1)和(2)所示。

$$R_{\text{sum}}(n) = \left| \sum_{i=0}^{K-1} R_i(n) \right|, n = 0, 1, \dots, L-1 \tag{1}$$

$$R_{\text{sum}}(n) = \sum_{n=0}^{K-1} |R_i(n)|, n = 0, 1, \dots, L-1 \tag{2}$$

其中 $R_i(n)$ 为第 i 个伪码周期内 C/A 码信号与本地码信号的相关值，周期长度为 L 。公式(1)实现了相邻 K 个伪码周期的相干积分。公式(2)实现了相邻 K 个伪码周期的非相干积分。

相干积分在信号累加的同时很好的抵消了噪声，从而提高信噪比；但是相干积分时间受到了导航比特跳变和多普勒频域的限制。非相干积分中的取模运算对信号和噪声同时进行了平方，增益累加效果不如相干积分。

2.2 相干积分和非相干积分增益

在分析处理增益之前，首先介绍信噪比(S/N) ^[2,3]。对于 GPS 接收机，表示 S/N 的方法按照带宽的不同有如下几种：以输入带宽为参考的输入信噪比 S/N、以 1kHz 为参考的输入信噪比 S/N 和以 1Hz 为参考的输入信噪比 C/N₀。通常噪声基底为-174 dBm/Hz 或者-114 dBm/MHz，根据输入功率电平即可得到信噪比。不同带宽的输入功率和 S/N 见表 1。

表 1 不同带宽的输入功率和 S/N

输入功率/dBm	输入 S/N/(dB/2MHz)	S/N/(dB/1kHz)	C/N ₀ /(dB/Hz)
-130	-19 dB	14 dB	44 dB
-140	-29 dB	4 dB	34 dB
-150	-39 dB	-6 dB	24 dB

常规条件下，带宽为 2MHz 的接收机输入端的 S/N 大约是-19 dB，这个信号太微弱以至于无法检测到，必须采取措施来提高信噪比。

1. 相干积分增益

由相应的带宽可以得到相干信号处理的增益，而带宽与数据长度相关^[2]。

如果对 1ms 的数据进行相干处理，等效带宽为 1kHz，噪声基底为-144 dBm。若输入为-130 dBm 的信号，则 1ms 的相干积分能产生的增益为：S/N=-130+144=14 dB。

如果对 2ms 的数据进行相干处理，等效带宽为 500Hz，噪声基底为-147 dBm。则 2ms 的

相干积分能产生的增益为： $S/N=-130+147=17\text{ dB}$ 。

一般来说，相对于 1ms ， $n\text{ ms}$ 的附加相干积分增益 G_c 为

$$G_c = 10\lg(n) \tag{3}$$

当 $n=2$ 和 10ms 时，附加增益分别为 $G_c=3$ 和 10dB 。由于导航数据每 20ms 可能出现一次相位跳变，捕获弱信号通常采用的最大长度是 10ms 。但是有时也用长于 10ms 的数据进行相干积分^[2]。图 1 给出了相干积分增益。

2. 非相干积分增益

非相干积分增益等于相干积分增益减去非相干积分损耗，而非相干积分损耗可由检查概率和虚警概率得到^[2]。

非相干积分损耗为：

$$L(n) = 10 \times \lg \left[\frac{1 + \sqrt{1 + 9.2n / D_c(1)}}{1 + \sqrt{1 + 9.2 / D_c(1)}} \right] \text{ dB} \tag{4}$$

式中， n 是非相干积分数目； $D_c(1)$ 为理想检测能力因子，是检测概率 P_d 和虚警概率 P_{fa} 的函数。一般取 $P_d = 0.9$ 和 $P_{fa} = 10^{-7}$ 得到 $D_c(1) \approx 21$ ^[3]。

则非相干积分增益 G_i 为：

$$G_i(n) = G_c(n) - L(n) = 10\lg(n) - L(n) \tag{5}$$

图 2 给出了非相干积分增益。

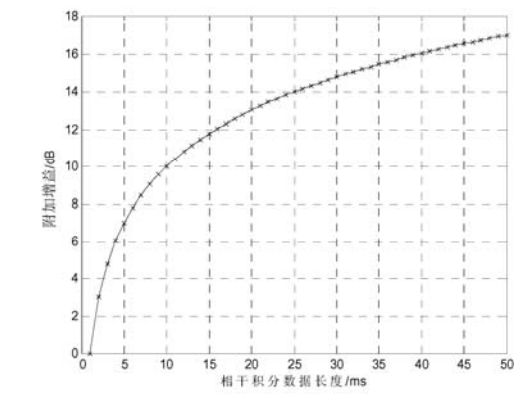


图 1 相干积分增益

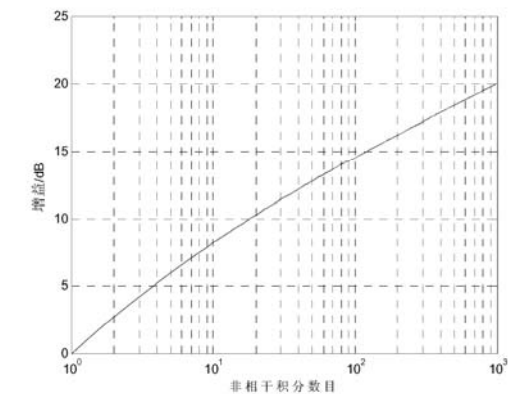


图 2 非相干积分增益

2.3 相干/非相干积分结合

雷达检测单个脉冲信号时常用的 S/N 是 14 dB 。而且由以往的捕获经验可以知道， $S/N=14\text{ dB}$ 时，能得到令人满意的捕获结果。下面就是通过采用相干和非相干积分结合的方法，达到 $S/N=14\text{ dB}$ 这一目标。

相干积分能得到较大的增益，但会受到数据跳变的影响,而非相干积分能缓解数据跳变的影响，但累加增益较小。为了得到更好的捕获效果，现在用得较多的累加方式是两者的结合，即根据数据位长度先进行相干积分后再进行非相干积分^[4]。假设每段相干积分数据长度为 $L(L$ 个 PRN 码周期)，非相干积分数目为 M ，则经过相干和非相干结合处理后的结果为：

$$P_{\text{long}} = \sum_{m=0}^{M-1} \left[\left(\sum_{l=l_m}^{l_m+L-1} I_l \right)^2 + \left(\sum_{l=l_m}^{l_m+L-1} Q_l \right)^2 \right] \tag{9}$$

式中，I 和 Q 分别为接收信号与本地载波 $\sin(2\pi f_L)$ 、 $\cos(2\pi f_L)$ 相乘，得到基带同向分量和正交分量。 $l_0, l_1, l_2, \dots, l_{M-1}$ 为每段相干积分时间的 PRN 码起始点。

选取不同的 L 和 M，就会获得不同的相干积分和非相干积分组合方式。下面介绍了一种方法，在给定输入信号强度情况下，确定合适的 L 和 M 来进行捕获。假定输入信号强度为-150 dBm(低于标准功率电平 20 dB)，具体步骤如下：

- 1) 确定相干积分时间长度。将 L 任取一个整数值，因为导航比特跳变的原因，一般 L 取 1~10 的整数。
 - 2) 计算相干积分增益 G_c 。对 Lms 的数据进行相干积分，相应的带宽为 $(1/L)\text{kHz}$ ，噪声基底为 $N=(-174+10\lg(1000/L))\text{dBm}$ 。则相干积分得到的增益为 $G_c=-150+N\text{ dB}$ 。
 - 3) 计算需要的非相干积分增益 G_i 。为了达到 $S/N=14\text{ dB}$ 这一目标，需要的非相干增益 $G_i=14-G_c$ 。
 - 4) 确定非相干积分次数。根据 G_i 的值，从图 2 中即可得到对应的 M 的值。
- 输入信号强度为-150 dBm，目标 $S/N=14\text{ dB}$ 情况下，一些 L 和 M 的组合如表 2 所示。

表 2 一些 L 和 M 的组合(输入信号强度为-150 dBm，目标 $S/N=14\text{ dB}$)

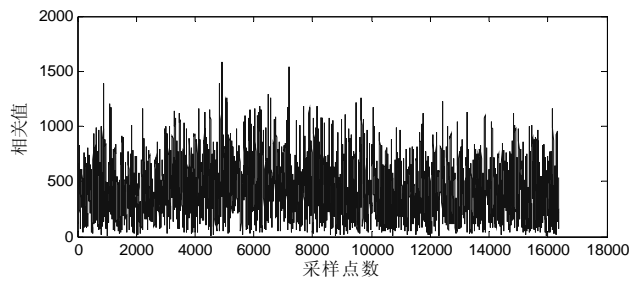
L	M	非相干积分时间/ms
1	1000	1000
2	273	546
4	80	320
8	26	208
10	19	190

在实际的捕获过程中，每 20ms 可能存在的导航比特相位跳变会导致积分增益的损失，考虑到这一点，那么要求的非相干积分会比表 2 中的结果稍长一些。

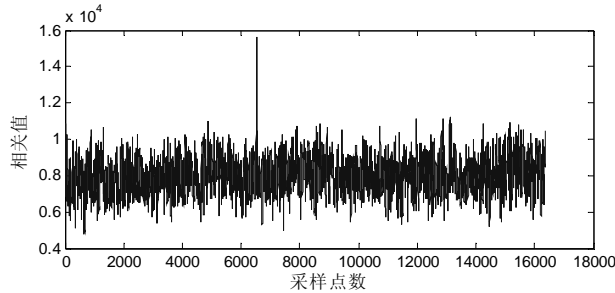
不同的相干/非相干积分组合方式都能捕获到信号，但是在计算量、频率分辨率存在上会存在差别。在实际数据的处理过程中，还要考虑到多普勒频率的影响。所以需要根据不同的信号选择不同的组合方式。

3 仿真结果

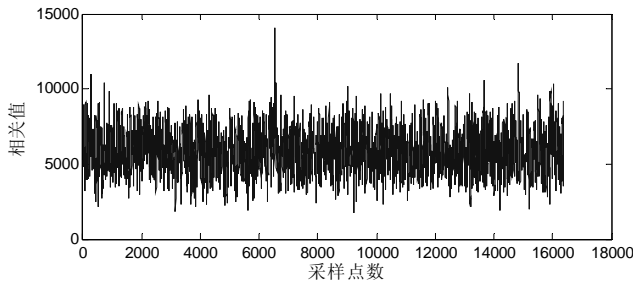
仿真过程中使用的数据为采集的真实 GPS 数据，对 11 号卫星进行捕获。已知该卫星 C/N_0 约为 37 dB，由 2.3 节计算得到，取 L=1 时，M=5；L=10 时，M=1。考虑到导航比特相位跳变带来的积分增益损失，要适当增加非相干积分次数 M 的值。最终选择 L=1，M=20 和 L=10，M=5 两种相干/非相干积分组合来进行捕获。图 3 中给出了三种不同情况下的捕获结果。



(a) $L=1, M=1$



(b) $L=1, M=20$



(c) $L=10, M=5$

图3 不同相干积分和非相干积分的捕获结果

图 3(a)即为 1ms 数据的相关捕获, 显然这种方法无法捕获弱信号。图 3(b)和 3(c)为所选择的两种相干/非相干组合的捕获结果, 捕获效果比 3(a)有明显改善。

基于 FFT 的相关运算, 操作数目与 $N \times \lg 2N$ 成正比^[2], 其中 N 是数据采样点数。实际 GPS 信号采样率为 16.3676MHz, 则(b) $L=1, M=20$ 时, 运算量为 1.478×10^6 ; (c) $L=10, M=5$ 时, 运算量为 4.513×10^6 。可以看出 3(c)的计算量远大于 3(b)。在不同的组合方式中, 一般选择运算量小的组合来进行捕获, 以缩短捕获时间。

4 结语

本文详细分析了相干积分和非相干积分在弱信号捕获过程中对信噪比的改善。但是当信号较弱时, 通常需要采用两种积分方式结合的方法来得到更高的信噪比。相干/非相干积分组合有多种组合方式, 本文给出了不同组合方式的确定方法。在实际捕获过程中, 如果已知信号强度, 就可以通过文中给出的方法, 确定合适的相干/非相干积分组合来进行捕获。

参 考 文 献

- [1] 蔡昌听,皮亦鸣,高灵敏度 GPS 技术的研究进展[J].全球定位系统,2006,31(2):1-4.
- [2] JAMES BAO-YEN TSUI. Fundamentals of global positioning system receivers : a software approach (2nd edition)[M]. John Wiley, 2005
- [3] BARTON,D.K., Modern radar system analysis[M]. Artech House, Norwood, MA, 1988
- [4] PSIAKI,M.L., Block acquisition of weak signals in a software receiver[C], ION GPS 2001,pp.2838-2850,Salt Lake City, UT, September 11-14 ,2001
- [5] 曹珂,王长林,梁宏, 低信噪比环境下短扩频码捕获性能的改善[J]. 电视技术,2008.48(7):69-73.

基于量子遗传神经网络的MIMO信号检测技术

周 敏 李 飞

(南京邮电大学 通信与信息工程学院, 江苏 南京 210003)

摘 要: MIMO 信号的最优检测在常规条件下是一 NP 难解问题, 利用 QGA (量子遗传算法) 和 RBF (径向基函数) 神经网络相结合, 提出了一种融合两种算法优点的 QGNN (量子遗传神经网络算法), 并将其应用于 MIMO 信号检测中。所提算法是把 QGA 的检测结果作为 RBF 神经网络的检测输入, 由于 QGA 给 RBF 网络提供了较好的初始值, 故能够使 RBF 网络快速收敛到最优解, 避免了由初始值的随机选取而带来的检测误码, 使检测性能得到进一步提高。仿真结果证明了该方法的检测性能优于传统检测器和一些使用智能算法的 MIMO 信号检测器。
关键词: 信息处理技术; MIMO 检测; 量子遗传神经网络; 量子遗传算法; RBF 神经网络

MIMO Detection Based on Quantum Genetic Neural Network

Zhou Min Li Fei

(College of Telecommunication&Information Engineering, NanJing University of Posts and Telecommunication, NanJing 210003,China)

Abstract: The optimal solution of MIMO signal detection is a NP(Non-deterministic Polynomial-time hard) problem. A Quantum Genetic Neural Network(QGNN) that employs a Quantum Genetic Algorithm(QGA) and a RBF (Radial Basis Function) neural network for the MIMO signals detection problem is proposed. Using this algorithm, the output of detector by the QGA as the input of detector by the RBF neural network to avoid the bit -error rate for selecting initial data randomly and improve further the detection property. The emulation results have shown that the proposed method has more powerful property than the conventional detector and some detectors based on intelligent algorithm.

Keywords: Information Processing Technology; MIMO Detection ; Quantum Genetic Neural Network ; Quantum Genetic Algorithm; RBF neural network

1 引言

最近二十多年以来, 移动通信在全球取得了巨大的发展, 并最终形成为与人类日常生活密切相关的产业。随着移动通信日益广泛的应用, 人们对其又提出了更高传输速率和更高传输质量的要求, 如手机上网、视频传输等等。在这种背景下, Bell 实验室于上个世纪九十年

代后期提出了一种新的通信系统结构,即 MIMO(Multiple-Input Multiple-Output)通信系统。该通信系统在其发射端和接收端均采用多个天线,从而实现了多个数据流在相同时间和相同频带的传输和接收。Foschini 和 Telatar 已经证明,MIMO 系统的信道容量随着发射天线数的增加呈近似线性的增长,这为 MIMO 通信方式在移动通信中的应用提供了坚实的理论基础。然而 MIMO 系统本身所提供的性能和增益能够有多少被挖掘出来,和接收机的算法有很大的关系,复杂的接收机检测算法也是影响 MIMO 系统大规模商用的一个原因,作为最优解码的最大似然检测(MLD)由于复杂度随着发射天线个数成指数性增加^[1],从而限制了它在实际系统中的应用。

与传统的单输入单输出(SISO)系统相比,MIMO 系统接收端接收到的是在时间上和频率上均相互重叠的多路信号,频率选择性系统中还存在不同时刻信号间的码间干扰,MIMO 信号检测面临着远高于传统 SISO 信号检测的困难和问题。由于信号检测性能的好坏将直接影响到整个 MIMO 系统性能的好坏,设计高性能、低复杂度的 MIMO 信号检测算法已成为 MIMO 通信中的一项具有重大意义的关键技术。

量子遗传算法(QGA-Quantum Genetic Algorithm)是量子计算与经典遗传算法(GA)相结合而产生的一个新的研究领域。算法利用了量子计算的量子并行、量子纠缠特性,采用了多状态基因量子比特编码方式和量子旋转门更新操作,引入动态和静态调整旋转角机制和量子变异,使得算法比经典遗传算法具有更强的并行处理能力、更快的收敛速度且比传统信号检测算法具有更高的效率。

神经网络具有信息分布式存储、大规模自适应并行处理和高度容错特性等优点,可应用于信号检测领域,RBF 神经网络是一种非线性信号处理技术,学习速度快,其网络结构具有自适应确定、输出与初始权值无关等优良特性。文献[2]用 RBF 神经网络得到了在 CDMA 系统环境下接近最优贝叶斯检测器的性能,文献[3]给出了垂直分层空时码的 MAP 检测器的 RBF 神经网络实现算法。可以研究将 QGA 与 RBF 网络相结合用于 MIMO 信号检测,取长补短,发挥两者的最大功效,基于这样的思路,本文设计的 MIMO 检测器将 QGA 的检测结果作为 RBF 神经网络的检测输入,由于 QGA 给 RBF 网络提供了较好的初始值,故能够使 RBF 网络快速收敛到最优解,避免了由初始值的随机选取而带来的检测误码,使检测性能得到进一步提高,计算机仿真结果表明该检测器能获得较好的检测性能。

2 MIMO接收机模型

考虑 M 根发射天线和 N 根接收天线的 MIMO 系统,假设为加性高斯白噪声(AWGN)信道,采用 QPSK 调制,输入输出关系可描述为:

$$y = Hx + n \tag{1}$$

式中 y ——接收天线收到的检测信号,维数: $N \times 1$;

x ——传输向量,维数: $M \times 1$;

H ——信道增益矩阵,维数: $M \times N$;

n ——零均值的高斯白噪声。

接收机的任务是从 $y = Hx + n$ 中检测出发射信号 x 。最大似然准则使发射为 x 的条件下接收到 y 的概率最大,即满足下式:

$$\max p_{y|x}(y|x) \quad (2)$$

由于已经假定信道噪声为加性高斯白噪声 (AWGN)，则在给定 H 和 x 下， y 的条件概率为：

$$p_{x|y}(y|H, x) = \frac{1}{(2\pi\delta_n^2)^{m/2}} \exp\left(-\frac{\|y - Hx\|^2}{2\delta_n^2}\right) \quad (3)$$

因此在给定 y 的条件下的最大似然估计为：

$$\begin{aligned} \hat{x} &= \arg \max_{x \in \mathbb{Z}^m} p_{x|y}(y|H, x) = \arg \max_{x \in \mathbb{Z}^m} \frac{1}{(2\pi\delta_n^2)^{m/2}} \exp\left(-\frac{\|y - Hx\|^2}{2\delta_n^2}\right) \\ &= \arg \min_{x \in \mathbb{Z}^m} \|y - Hx\|^2 \end{aligned} \quad (4)$$

根据 (4) 式寻求最优解的过程，其计算复杂度随发射天线个数的增加成指数增长，本质上是整数域的最小二乘问题，同时也是一个 NP (Non-deterministic Polynomial-time hard) 的问题，限制了它在实际系统中的应用。

3 量子遗传算法MIMO检测器

3.1 量子计算的基本概念

在量子算法中，最小的信息单位用量子位来表示，量子位有时称为量子比特，一个量子位不仅可以表示 0 和 1 两种状态，而且还可以同时表示这两种状态之间的任意叠加态，即一个量子位可以处于 $|1\rangle$ 或 $|0\rangle$ ，或者出于两者之间的中间态，即 $|1\rangle$ 和 $|0\rangle$ 的不同叠加态。因此一个量子位的状态可以表示为：

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle \quad (5)$$

其中， α 和 β 分别是 $|0\rangle$ 和 $|1\rangle$ 的概率幅，且满足下列归一化条件：

$$|\alpha|^2 + |\beta|^2 = 1 \quad (6)$$

在式 (6) 中， $|\alpha|^2$ 表示量子态的观测值为 0 的概率， $|\beta|^2$ 表示量子态的观测值为 1 的概率。满足式 (5) 和 (6) 的一对实数 α 和 β 称为一个量子位的概率幅，记为 $[\alpha, \beta]^T$ 。于是，任意一个有 m 个量子位的量子个体 P_j 可表示为：

$$P_j = \begin{bmatrix} \alpha_1 & \alpha_2 & \dots & \alpha_m \\ \beta_1 & \beta_2 & \dots & \beta_m \end{bmatrix} \quad (7)$$

其中， $|\alpha_i|^2 + |\beta_i|^2 = 1$, $i=1,2,3,\dots,m$ 。

3.2 基于量子遗传算法的MIMO检测器

为了实现 QGA 信号检测算法的良好性能，对 QGA 算法的主要参数进行如下设计[4]：

(1) 种群初始化：仿真时将量子染色体中量子比特基因数取系统的发射天线数 m ，所有

基因 α_i^t 和 β_i^t 初始化为 $1/2^{1/2}$ ，表示所有可能的叠加态以相同的概率出现；

(2) 适应度函数的选取：适应度函数是用来评估每个染色体的好坏，并且是非负的，根据最大似然准则：

$$\hat{x} = \arg \min_{x \in \Phi} \left\{ \|y - H \cdot x\|^2 \right\} = \arg \max_{x \in \Phi} \left[2x^T H^T y - x^T H^T H x \right] \quad (8)$$

取 MIMO 信号检测的目标函数为：

$$\Omega(x) = 2x^T H^T y - x^T H^T H x$$

由于 $\Omega(b)$ 值可正可负，为确保适应度函数的非负性，在 QGA 信号检测中，适应度函数设为：

$$f(b) = \exp(\mu \cdot \Omega(b))$$

式中 μ 为一正常数，仿真时取 0.05。这样 MIMO 信号最佳检测就被描述为一个 QGA 求最优个体的问题。

4 量子遗传神经网络MIMO检测器

4.1 基于RBF神经网络的MIMO信号检测

典型的径向基函数是高斯函数：

$$\varphi_i(y) = \exp \left[-\frac{\|y - c_i\|^2}{\delta_i^2} \right], i = 0, 1, 2 \dots n_c \quad (9)$$

其中 y 为输入变量， c_i 为中心， δ_i 代表了高斯函数的宽度， n_c 为网络隐层结点数。径向基函数神经网络的输出响应可表示为：

$$F(y) = \sum_{i=0}^{n_c} w_i \varphi_i(y) \quad (10)$$

当神经网络的输入结点数与信号检测器的输出数 n_r 相等时，可以将接受机的输出信号作为网络的输入信号。对网络进行如下设置： $w_0=0$ ， $n_c = 2^{n_r}$ ，径向基函数取(9)式所示的高斯函数，中心 c_i 取 n_c 种对应于不同发射符号矢量而产生的接收符号矢量的无噪分量， $y\% = Hx$ ， $\delta_i = 1$ 中心 c_i 对应的发射符号矢量的径向基函数的权重 $w_i = 1$ ，剩余的 $w_i = -1$ 。经过以上设置之后，径向基函数神经网络就等价于一个最大似然判决检测器。基于 RBF 神经网络的 MIMO 信号检测算法流程如下[5]：

Step1: 选择一组训练样本，采用基于最近邻聚类中心选取的 RBF 神经网络学习算法确定隐层各结点的中心值、高斯宽度、和网络连接权值；

Step2: 根据 (9) 式所示的径向基函数计算隐层节点的结果；

Step3: 计算神经网络的输出： $F = \sum_{n=1}^{n_c} w_n \varphi_n$ ；

Step4: 根据 $\hat{x}_i = \text{sgn}(F(y)) = \begin{cases} 1, & F(y) \geq 0 \\ -1, & F(y) < 0 \end{cases}$ 对 \hat{x}_i 进行判决。 $i=i+1$ 。

回到 Step1，直到将所有的发射符号分量都判决完，此时完成一个时间间隔接受到的符号矢量的判决，检测下一个时间间隔接收到的符号矢量并从第一步重新开始。

4.2 基于量子遗传神经网络的MMO信号检测

设计 MIMO 检测器将 QGA 的检测结果作为 RBF 神经网络的检测输入，由于 QGA 给 RBF 网络提供了较好的初始值，故能够使 RBF 网络快速收敛到最优解，避免了由初始值的随机选取而带来的检测误码。

该检测器中的径向基函数神经网络采用 3 层网络结构，分别是输入层、隐层和输出层，量子遗传神经网络 MIMO 信号检测器结构如图 1 所示：

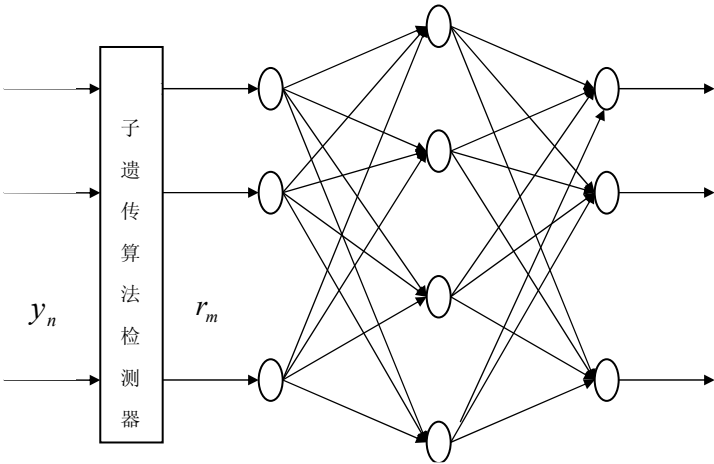


图 1 量子遗传神经网络 MIMO 信号检测器

图中 QGA 检测器有 N 个输入和 M 个输出单元，QGA 接收 MIMO 接收机的输出： $y = [y_1, y_2, \dots, y_n]^T$ (n 为接收天线数)，经过多次量子交叉和变异操作，得到一个收敛解。然后该收敛解被作为 RBF 神经网络检测器的初始输入： $r = [r_1, r_2, \dots, r_m]^T$ (m 为发射天线数)，网络实时快速优化计算，得到发射信号的检测输出： $\hat{x} = [\hat{x}_1, \hat{x}_2, \dots, \hat{x}_m]^T$ 。

5 仿真实验及结果分析

根据以上提到的模型，使用 Matlab 仿真软件，MIMO 系统假设基于多天线瑞利衰落信道，采用 QPSK 调制，噪声为均值为 0 的独立同分布高斯白噪声，我们分别考虑 4 个发送、4 个接收天线和 8 个发送、8 个接收天线，并且假设信道矩阵 H 在每 T=1000 个符号周期内保持不变，然后随机的改变，我们考虑了以下检测器：（1）传统检测器（最小均方误差 MMSE）；（2）基于经典遗传算法的 MIMO 信号检测器（GA）；（3）基于量子遗传算法的 MIMO 信号检测器（QGA）；（4）基于 RBF 神经网络的 MIMO 信号检测器（RBF）；（5）基于量子遗传神经网络

的 MIMO 信号检测器 (QGNN)。实验结果表明在 4 发 4 收和 8 发 8 收的条件下, 本文提出的 QGNN 检测器都比其他检测器获得了更好的检测效果, 误码率达到更低:

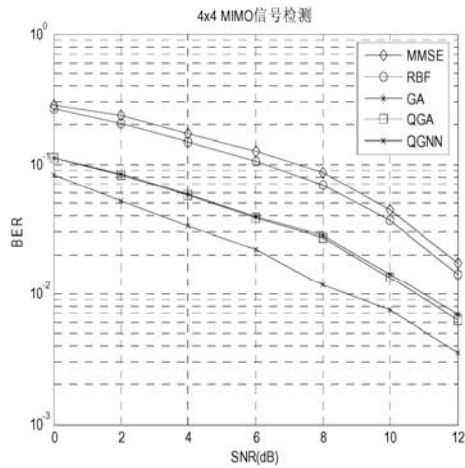


图 2 4 发 4 收时算法性能比较

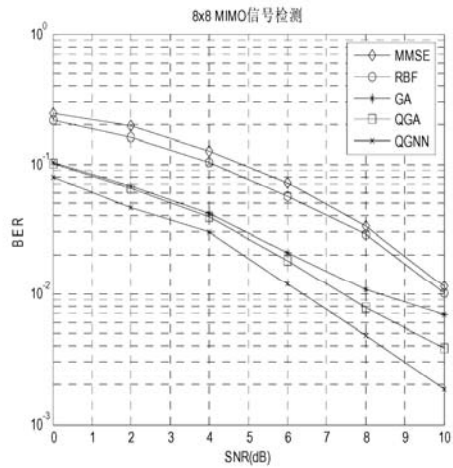


图 3 8 发 8 收时算法性能比较

6 结论

本文提出了一种将量子遗传算法与 RBF 神经网络相结合的量子遗传神经网络 (QGNN) 算法用于 MIMO 信号检测的方案, 综合利用了量子计算的量子并行计算、量子纠缠特性和遗传算法较好的全局收敛性能以及 RBF 神经网络的大规模自适应并行处理、快速学习能力, 使 QGNN 比传统算法具有更强的并行处理能力和更快的收敛速度。计算机仿真结果表明: 本文提出的检测器方案比传统检测器和一些基于智能算法的检测器有更良好的检测性能。

参 考 文 献

[1] Daniel K.C.So,Roger S.Cheng. Layered maximum likelihood detection for MIMO systems in frequency selective fading channels[A]. IEEE TRANSATIONS ON WIRELESS COMMUNICATIONS,VOL 5,NO.4,APRIL 2006

[2] CHEN S,SAMINGAN A K,HANZO L.Adaptive nears minimum error rate training for neural networks with application to multiuser detection in CDMA communication systems[J].Signal Processing,2005,85:1435-1448.

[3] 陈其铭, 尹俊勋, 垂直分层空时码的 MAP 检测器及其神经网络实现[J], 电子与信息学报, 2004 年 9 月, 第 26 卷第 9 期.

[4] 洪亮, 李飞, 基于量子遗传算法的 CDMA 多用户检测技术[J], 信息技术, 2008.03

[5] 杨倩倩, 基于量子神经网络的 MIMO 信号检测技术的研究[D], 南京邮电大学, pp:32-34.

作者简介:

周敏 (1985-), 男, 江苏常州, 南京邮电大学通信与信息工程学院硕士研究生, 主要研究方向: 量子信息处理技术;

李飞 (1966-), 女, 湖南长沙, 南京邮电大学通信与信息工程学院教授, 主要研究方向: 现代通信中的智能信号处理技术、量子信息处理技术。

基于最大似然聚类的GMM优化方法 及其在说话人辨认中的应用

胡 婕 周 琳

(东南大学 信息科学与工程学院, 江苏 南京 210096)

摘 要: 模式识别中基于高斯混合模型 GMM(Gaussian Mixed Model)的说话人辨认系统在训练样本充分的条件下获得了较高的识别率, 但其计算复杂度往往限制了系统应用。从提高系统实用性的角度, 介绍了一种基于最大似然 ML(Maximum Likelihood)聚类的简化 GMM 算法。仿真结果表明在保证系统识别性能的前提下, 简化算法有效降低了计算开销。

关键词: 模式识别; 高斯混合模型; 最大似然; 说话人辨认

GMM Optimization Based on ML Clustering and Its Application in Speaker Identification

Hu Jie Zhou Lin

(Department of Information Science and Engineering, Southeast University, Nanjing 210096, China)

Abstract: In the domain of pattern recognition, although speaker identification system based on GMM (Gaussian Mixed Model) has achieved high performance with sufficient training data, its application is restricted by a large complexity of computation. This paper presents a simplified GMM algorithm based on ML (Maximum Likelihood) cluster for practical application. The simulation result indicates that the method reduces the computational complexity, while keeps a high identification performance.

Keywords: pattern recognition; GMM; ML; speaker identification

说话人辨认是根据语音信号中的个人特征决定说话人身份的过程。在与文本无关的辨认方式下, 高斯混合模型以及由其衍生出来的统计学模型是目前较为有效的方法。通常, GMM 方法对所有用户使用统一的模型结构, 通过期望最大化 EM(Expectation Maximization)算法进行参数估计, 从而建立说话人的概率模型。然而, EM 算法的计算复杂度较大、收敛速度慢, 使得传统 GMM 算法很难满足大数据集应用中的实时性要求。

在实际应用中, 运算速度和识别率是同时需要考虑的问题。识别率要求一定数量的训练语音样本, 而运算速度则随训练样本数的增加而减慢。在使用 DSP 实时实现说话人辨认系统时, 算法计算量是需要解决的关键问题。

1 基于GMM的说话人辨认系统

混合阶数为 M 的 GMM 可以表示为：

$$p(\mathbf{x}|\lambda) = \sum_{i=1}^M w_i \xi_i(\mathbf{x}) \quad (1)$$

其中, \mathbf{x} 为 D 维观测矢量, $\xi_i(\mathbf{x})$ 为 GMM 的第 i 个高斯分量, w_i 为混合权重, 满足 $\sum_{i=1}^M w_i = 1$ 。

D 维高斯函数可以表示为：

$$\xi_i(\mathbf{x}) = \frac{1}{(2\pi)^{D/2} |\Sigma_i|^{1/2}} \exp \left\{ -\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu}_i)' \Sigma_i^{-1} (\mathbf{x} - \boldsymbol{\mu}_i) \right\} \quad (2)$$

其中, $\boldsymbol{\mu}_i$ 为均值矢量, Σ_i 为协方差矩阵。整个 GMM 由各混合分量的均值矢量、协方差矩阵及混合权重描述, 可用模型参数 $\lambda = \{w_i, \boldsymbol{\mu}_i, \Sigma_i, i = 1, \dots, M\}$ 表示。

基于最大似然 (ML) 准则的 GMM 参数估计问题可以描述为：

$$\hat{\lambda}_{\text{ML}} = \arg \max_{\lambda} p(\mathbf{X} | \lambda) \quad (3)$$

其中, 观测矢量序列 $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N\}$ 的似然函数为：

$$p(\mathbf{X} | \lambda) = \prod_{n=1}^N p(\mathbf{x}_n | \lambda) \quad (4)$$

式 (3) 的最大化过程可以通过 EM 算法来进行^[1]。该算法迭代求解 GMM 参数, 每步迭代可以使估计模型的似然函数值单调递增。当前后两次迭代所得的似然函数值之差小于预先设定的门限时, 即输出估计结果。实验表明, GMM 的协方差阵可以用高阶对角阵近似, 这种近似在保证识别性能的同时可有效降低系统运算量^[2]。此时, 协方差阵 Σ_i 的估计简化为方差 σ_i 的估计。可以证明^[1], 下面的重估公式保证模型收敛到局部最优。

$$\hat{w}_i = \frac{1}{N} \sum_{n=1}^N p(i | \mathbf{x}_n, \lambda) \quad (5)$$

$$\hat{\boldsymbol{\mu}}_i = \frac{\sum_{n=1}^N p(i | \mathbf{x}_n, \lambda) \mathbf{x}_n}{\sum_{n=1}^N p(i | \mathbf{x}_n, \lambda)} \quad (6)$$

$$\hat{\sigma}_i = \frac{\sum_{n=1}^N p(i | \mathbf{x}_n, \lambda) \mathbf{x}_n^2}{\sum_{n=1}^N p(i | \mathbf{x}_n, \lambda)} - \boldsymbol{\mu}_i^2 \quad (7)$$

其中, 后验概率为：

$$p(i | \mathbf{x}_n, \lambda) = \frac{w_i \xi_i(\mathbf{x}_n)}{\sum_{j=1}^M w_j \xi_j(\mathbf{x}_n)} \quad (8)$$

在说话人辨认系统中，每个说话人可以由一个参数为 $\lambda^{(k)}$ ($k = 1, \dots, K, K$ 为闭集人数) 的 GMM 代表。辨认的任务是找到一个说话者 k^* ，其对应的模型参数 $\lambda^{(k^*)}$ 使得待识别语音的特征矢量序列 $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N\}$ 具有最大似然度，即：

$$k^* = \arg \max_{1 \leq k \leq K} p(\mathbf{X} | \lambda^{(k)}) = \arg \max_{1 \leq k \leq K} \ln p(\mathbf{X} | \lambda^{(k)}) \quad (9)$$

2 改进GMM算法

式 (1) 中，观测矢量 \mathbf{x}_n 的似然函数是由 M 个高斯概率密度加权求和得到的，该步计算可以通过维特比算法简化。维特比算法解决了给定观测矢量序列 $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N\}$ 以及模型参数 λ 时，如何在最大似然的意义上确定一个最佳状态序列 $S = \{s_1, s_2, \dots, s_N\}$ 的问题。

实验表明^[3]，在 GMM 中，对于单个观测矢量 \mathbf{x}_n ，只有一个高斯分量 i_n 对应的概率值较大，即单个观测矢量落入混合模型中某个高斯分布的可能性较大，因此其似然函数可以用该分布的概率密度函数 $\xi_{i_n}(\mathbf{x}_n)$ 近似，而 i_n 即代表了观测矢量所对应的状态。 \mathbf{x}_n 对应的高斯分量可以通过下式求解：

$$\begin{aligned} i_n &= \arg \max_{1 \leq i \leq M} \xi_i(\mathbf{x}_n) \\ &= \arg \max_{1 \leq i \leq M} \{\ln \xi_i(\mathbf{x}_n)\} \\ &= \arg \min_{1 \leq i \leq M} \left\{ (\mathbf{x}_n - \boldsymbol{\mu}_i)' \boldsymbol{\Sigma}_i^{-1} (\mathbf{x}_n - \boldsymbol{\mu}_i) + \ln |\boldsymbol{\Sigma}_i| \right\} \end{aligned} \quad (10)$$

此时，观测矢量 \mathbf{x}_n 的似然函数简化为：

$$p(\mathbf{x}_n | \lambda) \approx \xi_{i_n}(\mathbf{x}_n) \quad (11)$$

第 i 分量的后验概率简化为：

$$p(i | \mathbf{x}_n, \lambda) \approx \begin{cases} 1 & i = i_n \\ 0 & i \neq i_n \end{cases} \quad (12)$$

将式(11)和式(12)代入式(4)至(9)得改进算法流程如下。

训练流程：

- 1) 输入训练样本序列 $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N\}$ ，设置最大迭代次数为 MaxIter；
- 2) 初始化^[4]：从训练样本集中均匀抽取 M 个样本作为均值向量的初始值 $\boldsymbol{\mu}_i^{(0)}$
协方差阵 $\boldsymbol{\Sigma}_i^{(0)} = \mathbf{I}$
迭代次数 $m = 0$
- 3) 找到每个样本点 \mathbf{x}_n 对应的高斯分量：

$$i_n = \arg \min_{1 \leq i \leq M} \left\{ (\mathbf{x}_n - \boldsymbol{\mu}_i^{(m)})' (\boldsymbol{\Sigma}_i^{(m)})^{-1} (\mathbf{x}_n - \boldsymbol{\mu}_i^{(m)}) + \ln |\boldsymbol{\Sigma}_i^{(m)}| \right\} \quad (13)$$

- 4) 定义 $S_i = \{\mathbf{x}_n \in \mathbf{X} | i_n = i\}$ 为每个状态对应的样本集，更新模型参数：

$$\boldsymbol{\mu}_i^{(m+1)} = \frac{1}{N_i} \sum_{\mathbf{x}_n \in S_i} \mathbf{x}_n \quad (14)$$

$$\boldsymbol{\sigma}_i^{(m+1)} = \frac{1}{N_i} \sum_{\mathbf{x}_n \in S_i} \mathbf{x}_n^2 - (\boldsymbol{\mu}_i^{(m+1)})^2 \quad (15)$$

其中 N_i 为 S_i 中包含的样本点数;

- 1) 判断 $m < \text{MaxIter}$? 若否, 转入 6) 执行, 否则, 令 $m = m + 1$, 转入 2) 执行;
- 2) 迭代终止, 输出模型参数:

$$\begin{aligned} w_i &= \frac{N_i}{N} \\ \mu_i &= \mu_i^{(m)} \\ \Sigma_i &= \text{diag}\{\sigma_i^{(m)}\} \end{aligned}$$

识别流程:

- 1) 输入待识别的特征序列 $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N\}$;
- 2) 找到样本 \mathbf{x}_n 在模型 $\lambda^{(k)}$ 下对应的高斯分量:

$$i_n^{(k)} = \arg \min_{1 \leq i \leq M} \left\{ (\mathbf{x}_n - \mu_i^{(k)})' (\Sigma_i^{(k)})^{-1} (\mathbf{x}_n - \mu_i^{(k)}) + \ln |\Sigma_i^{(k)}| \right\} \quad (16)$$

则该模型下, 特征矢量序列的似然函数:

$$p(\mathbf{X} | \lambda^{(k)}) \approx \prod_{n=1}^N \xi_{i_n}^{(k)}(\mathbf{x}_n) \quad (17)$$

其中 $k = 1, \dots, K$ 为对应的说话人;

- 3) 输出识别结果:

$$\begin{aligned} k^* &= \arg \max_{1 \leq k \leq K} \left\{ \sum_{n=1}^N \ln \xi_{i_n}^{(k)}(\mathbf{x}_n) \right\} \\ &= \arg \min_{1 \leq k \leq K} \left\{ \sum_{n=1}^N \left[(\mathbf{x}_n - \mu_{i_n}^{(k)})' (\Sigma_{i_n}^{(k)})^{-1} (\mathbf{x}_n - \mu_{i_n}^{(k)}) + \ln |\Sigma_{i_n}^{(k)}| \right] \right\} \end{aligned} \quad (18)$$

上述改进算法避免了高斯分布函数中的指数和除法运算, 可通过乘累加和比较逻辑实现, 在 DSP 上实现时, 可有效提高识别系统的实时性。

3 实验结果

3.1 语音库描述

实验采用 TIMIT 语音库, 该语音库为纯净的阅读语音, 16kHz 采样, 16bit 线性量化。实验中选取了 100 名说话人的“sx”语音, 该类语音对于所有说话人的语料内容相同。每个人 5 句话, 每句话大约持续 3 秒, 其中 3 句用于训练, 2 句用于识别。由于整句话较短, 因此没有采用分段测试, 而是采用整句测试。整句测试时识别率的计算, 即用所有正确识别的人数除以总人数。

3.2 特征提取

语音特征采用 12 维 MFCC 参数, 特征提取的其它参数设置为:

- 1) 语音信号 16kHz 采样, 帧长 256 个采样点, 帧移 128 个采样点;
- 2) 预加重系数 $\alpha = 0.97$;
- 3) 数据窗采用 Hamming 窗, 256 点 FFT;

- 4) Mel 滤波器的个数为 24;
- 5) 由于使用的是 GMM 模型, 因此不对特征进行倒谱加权。

3.3 收敛性能

图 1 比较了原始 GMM 算法和改进算法的收敛性能。其中, 两种算法的模型阶数均取为 16, 纵坐标为所有训练语音帧对数似然函数求和所得。可以看出, 改进算法的收敛性能较原始算法有所降低, 但仍可保证训练模型以较快速度收敛到局部最优。

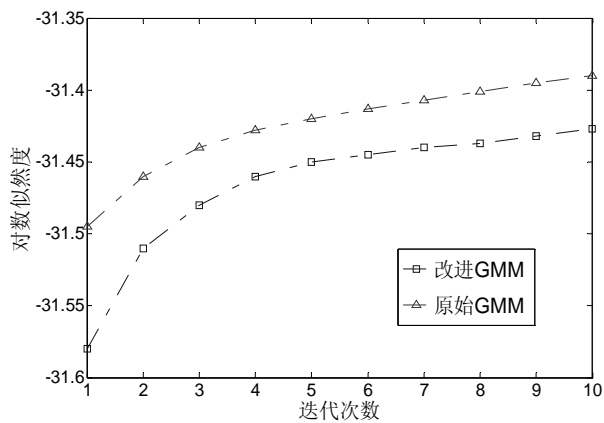


图 1 收敛性能比较

3.4 识别性能

表 1 比较了原始 GMM 算法和改进算法的 CPU 运行时间。该仿真在 CPU 主频为 2.66GHz 的 Pentium(R) 4 机器上进行。其中, 两种算法的迭代次数均设为 13 次, 训练和识别时间均为一个人的平均时间。可以看出, 在相同模型阶数条件下, 改进算法提高了系统的运行速度。

由改进算法的训练、识别流程可以看出, 计算开销的节省主要体现在以下几个方面:

- 1) 简化算法计算每个特征向量的似然函数时, 用单模高斯模型近似 GMM, 这种近似避免了式 (1) 中的 M 次乘累加操作;
- 2) 简化算法不需要计算似然函数的具体值, 只需要比较对数似然函数的大小, 从而避免了式 (2) 中的指数运算;
- 3) 简化算法的模型更新公式只需乘上相应的比例因子, 避免了原始公式 (6) (7) 中的除法操作;
- 4) 简化算法不需要计算式 (8) 中的后验概率, 只需根据比较逻辑的结果, 对特征向量的后验概率进行 2 进制赋值操作。

表 1 CPU 时间比较 (单位: 秒)

		改进 GMM	原始 GMM
M = 16	训练	0.33	0.45
	识别	2.88	3.02
M = 32	训练	0.65	0.75
	识别	5.62	6.13
M = 64	训练	1.33	1.47
	识别	11.71	12.64

表 2 比较了原始 GMM 算法和改进算法的识别率性能，实验条件同上。该实验共进行了 3 次，每次从语音库中抽取 100 个说话人进行测试，相应的识别率为 3 次的平均值。可以看出，改进算法在提高系统运行速度的同时并未显著降低识别率性能。但该算法对模型阶数的选择更为敏感，当模型阶数选择不合适时，会对识别率有较大影响。

表 2 识别率比较（单位：%）

	改进 GMM	原始 GMM
M = 16	88.7	91.0
M = 32	87.3	89.7
M = 64	86.3	87.0

3.5 模型阶数与迭代次数的选择

图 2 为模型阶数 M = 16、32、64 时识别率与迭代次数的关系曲线。所有实验均在相同的数据集中进行。由该图可知，对于实验中采用的语音库，当模型阶数取为 32 时可获得较好的识别性能。识别率随迭代次数递增，且当迭代次数大于 10 次时，模型趋于稳定。M = 64 时，模型出现欠拟合，单纯增加迭代次数对识别性能的提升影响不大。

由于改进算法对 GMM 参数估计问题进行了简化，每个特征向量的似然函数值仅用单模高斯模型近似，当模型阶数较小时，不能充分拟合特征空间的分布，增加了模型参数的估计误差；同时，GMM 参数的个数随模型阶数呈线性递增，当模型阶数取的较大时，对应于较大的参数集，参数估计误差对识别性能的影响也较大，但这种影响可以通过增加训练的数据量克服。

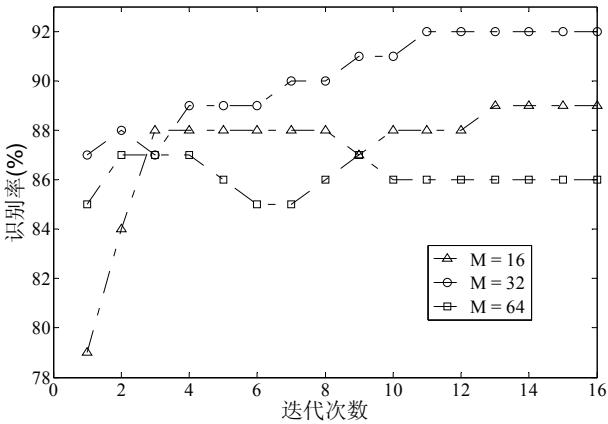


图 2 不同模型阶数的识别率与迭代次数关系曲线

4 结束语

本文从基于 GMM 说话人辨认系统的实用性角度出发，介绍了一种基于最大似然聚类的简化算法，并将该算法与原始 GMM 算法进行了比较。改进算法避免了计算高斯分布函数时的指数和除法运算，更适合 DSP 实时实现。实验表明，该算法在提高系统运行速度的同时不

会显著降低识别率性能，但对模型阶数的选择较为敏感，通常将模型阶数取为保证识别率性能的最小值，以节省系统开销。

参 考 文 献

- [1] BISHOP C.M., JORDAN M., KLEINBERG J., Pattern Recognition and Machine Learning [M]. New York: Springer-Verlag, 2006.
- [2] REYNOLDS D.A., DODDINGTON G, DUNN R.B., Automatic speaker recognition using Gaussian mixture speaker models [J]. Lincoln Lab, 1996, J. 8: 173-192.
- [3] HAUTAMAKI V., KINNUNEN T., KARKKAINEN I., Maximum a Posteriori Adaptation of the Centroid Model for Speaker Verification [J]. IEEE Signal Processing Letters, 2008, Vol. 15: 162-165.
- [4] REYNOLDS D.A., ROSE R.C., Robust Text-Independent Speaker Identification Using Gaussian Mixture Speaker Models [J]. IEEE Trans. on Speech and Audio Processing, 1995, 3(1): 72-83.

作者简介：

胡婕（1983-），女，江苏徐州，硕士研究生，东南大学，研究方向为语音信号处理；
周琳（1978-），女，博士，江苏镇江，东南大学，副教授，研究方向为语音信号处理。

一种新的基于段长分布的语音识别模型

杜树木 何良华

(同济大学 计算机科学与技术系, 上海 201804)

摘要: 在语音识别领域, 隐马尔可夫模型(Hidden Markov Model, 简称 HMM)是一个重要的方法。然而在其模型构建中, 状态的段长概率随时间呈指数下降, 没有恰当地表示语音的时序结构。本文基于动态贝叶斯网络(Dynamic Bayesian Network, 简称 DBN)的通用算法, 提出了在 DBN 模型中显式表示段长的改进算法。Aurora-5 语音库实验表明: 基于段长分布的 DBN 模型识别率在噪音环境下优于传统的 HMM 模型, 尤其在插入类型的错误方面。

关键字: 语音识别; 动态贝叶斯网络; 段长; GMTK

Novel Speech Recognition Model Based on Duration Distribution

Du Shu mu He Liang hua

(Department of Computer Science and Technology, Tongji University, Shanghai 201804, China)

Abstract: Hidden Markov Model (HMM) is an important approach in speech recognition. In model construction, probability of duration of a state decreases exponentially with time which is not appropriate for representing the temporal structure of speech. Although some improvements to duration models were proposed, they required significant modifications to training and decoding algorithms. Based on general algorithms of Dynamic Bayesian Network (DBN) model, this paper investigates how to explicitly represent duration in DBN model. Experiments on Aurora-5 database show that in noisy environments recognition rates of duration-based DBN model are better than those of traditional HMM, especially in insertion error.

Keywords: speech recognition; dynamic Bayesian network; duration; GMTK

1 前言

当前对自动语音识别(Automatic Speech Recognition, 简称 ASR)的研究主要基于音段特征如 MFCC, PLP 等, 这些特征在无噪音环境下识别效果很好, 但在噪音环境下识别率有明显的下降。而段长、声调等韵律特征在噪声环境下相对健壮^[1], 因此对韵律特征建模并加入到语音识别模型中以提高系统的鲁棒性和识别率, 逐渐成为当前语音识别的研究热点之一。

目前主流的语音识别系统利用隐马尔可夫模型(Hidden Markov Model, 简称 HMM)来模拟生成语音信号的语音单元。在 HMM 中, 各状态的段长分布隐式地包含在指向状态自身的弧中, 其有如下的几何分布:

$$p(d) = a_{ii}^{d-1}(1 - a_{ii}) \tag{1}$$

其中 a_{ii} 为状态 i 到自身的转移概率, d 为在状态 i 的驻留长度, 即在状态 i 对应的观察特征个数。在此分布中, 段长呈指数下降, 这被认为是 HMM 的一个缺陷^[2]。针对此缺陷, 很多学者对 HMM 的段长模型进行了改进^[3-4]。但这些改进并没有完全脱离 HMM 这个框架, 都是对 HMM 的一些假设进行了修改或放宽, 破坏了 HMM 高效算法所依赖的条件, 同时修改后的已经很复杂的模型较难再与其它韵律特征结合在一起。

解决这些问题的一个可行方法就是引入更一般、表达能力更强的模型, 以更自然地融合段长模型。动态贝叶斯网络(Dynamic Bayesian Network, 简称 DBN)^[5]模型作为 HMM 的扩展和推广, 非常适合处理语音这类复杂的且具有很强时序性的数据集。自从 Zweig^[6]将 DBN 用于一个实际的语音识别系统中, DBN 在语音识别领域得到越来越多的关注。特别是 Bilmes^[7]等人开发出针对 DBN 的图模型工具包(Graphical Models Toolkit, 简称 GMTK)后, 很多与语音相关的韵律特征被考虑进模型中来, 改善了识别效果。在 DBN 中, 每考虑一个新的段长模型, 很少涉及对基本算法的改动, 利用其通用的算法即可, 且各种段长模型可以在统一的框架下进行比较。

2 DBN简介

一个贝叶斯网络(BN)模型是一个概率图模型, 其通过有向无环图来表示一组随机变量及其条件独立关系。DBN 是对 BN 在时序过程建模方面的扩展, 是对复杂随机过程的一个简洁表示。一个 DBN^[8]由一个初始网 B_0 和转化网 B_{-} 组成, 即 (B_0, B_{-}) 。其中 B_0 是一个 BN, 其指定随机过程的初始状态分布 $P_{B_0}(Z_0)$ 。转换网 B_{-} 含有两个 BN, 其指定从时间片 $t-1$ 到时间片 t 的状态转换分布 $P_{B_{-}}(Z_t | Z_{t-1})$, 其计算公式为:

$$P_{B_{-}}(Z_t | Z_{t-1}) = \prod_{i=1}^N P_{B_{-}}(Z_t^{(i)} | Pa(Z_t^{(i)})) \quad (2)$$

其中 N 表示每个时间片上的结点数, $Z_t^{(i)}$ 表示时间片 t 上的第 i 个结点, $Pa(Z_t^{(i)})$ 表示 $Z_t^{(i)}$ 的父结点。图 1(a)给出了一个简单的 DBN 例子。由 (B_0, B_{-}) 定义的 DBN 对应一个在 Z_0, Z_1, Z_2, \dots 变量集上半无穷网络。实际上, 一般只在有穷时间间隔 $0, 1, \dots, T$ 上推理, 从而将一个 DBN 展开成在 Z_0, Z_1, \dots, Z_T 上的“长”贝叶斯网络, 其联合分布如下:

$$\begin{aligned} P(Z_{0:T}) &= P_{B_0}(Z_0) \prod_{t=1}^T P_{B_{-}}(Z_t | Z_{t-1}) \\ &= \prod_{i=1}^N P_{B_0}(Z_0^{(i)} | Pa(Z_0^{(i)})) \times \prod_{t=1}^T \prod_{i=1}^N P_{B_{-}}(Z_t^{(i)} | Pa(Z_t^{(i)})) \end{aligned} \quad (3)$$

图 1(b)给出了图 1(a)所示的 DBN 展开 4 个时间片相应的贝叶斯网络。在此例子中, $Z_t^{(1)} = X_t, Z_t^{(2)} = Y_t, Pa(X_t) = X_{t-1}, Pa(Y_t) = X_t$, 从而(3)式可写为:

$$P(X_{0:3}, Y_{0:3}) = P(X_0)P(Y_0 | X_0) \times \prod_{t=1}^3 P(X_t | X_{t-1})P(Y_t | X_t) \quad (4)$$

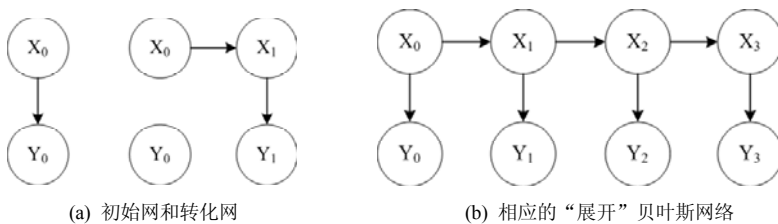


图 1 一个动态贝叶斯网络结构图

“动态”指模拟的是一个动态系统，而不是模型的结构随时间动态的变化。一般也假设时间片间的参数是不变的，从而使得 DBN 可以用有限的参数来模拟任意多的数据。

3 基于段长的DBN模型

在 HMM 中，每一个时间片只有单个变量用于表示状态，段长分布被隐式地表示在指向状态自身的弧中，如图 2 所示。如前言中所述，此段长为几何分布。而在 DBN 中，通过删除指向状态自身的弧来去掉此隐式段长分布，且显式地加入一个表示段长分布的变量，如图 3 所示。从而模型对下一个状态的预测不仅取决于前一个状态，还取决于在前一个状态已经驻留的时间。

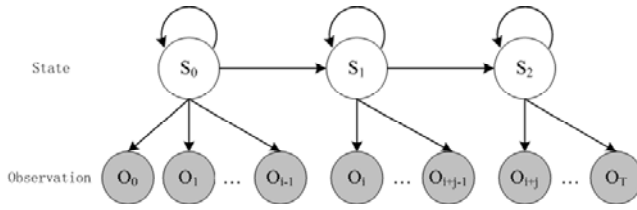


图 2 一个隐马尔可夫模型

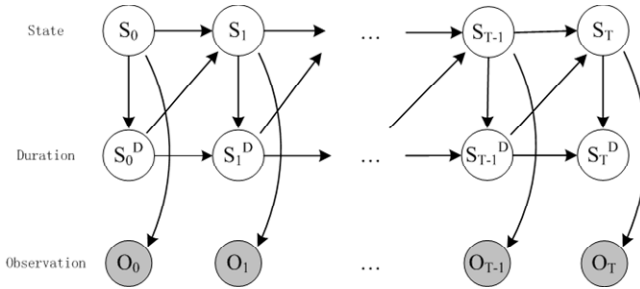


图 3 一个展开的含显式段长变量的动态贝叶斯网络模型

用 $P_i(d)$ 表示在状态 i 驻留时间为 d 的概率分布，当第一次进入状态 i 时，利用 $P_i(d)$ 的分布为 S_t^D 选择一个值，时间片每向前进一步， S_t^D 的值减 1，直到 S_t^D 为 0 时，此时状态才可以转移，利用状态转移概率分布 $A(i, j)$ 选择下一个状态 j ，针对此状态 j ，利用 $P_j(d)$ 再选择一个段长值，以此类推，直到结束。

对应图 3，式(3)中的 $P_{B_0}(Z_0^{(i)} | Pa(Z_0^{(i)}))$ 和 $P_{B_{\rightarrow}}(Z_t^{(i)} | Pa(Z_t^{(i)}))$ 可分别写为：

$$P_{B_0}(Z_0^{(i)} | Pa(Z_0^{(i)})) = P(S_0)P(O_0 | S_0)P(S_0^D | S_0) \quad (5)$$

$$P_{B_{\rightarrow}}(Z_t^{(i)} | Pa(Z_t^{(i)})) = P(O_t | S_t)P(S_t^D | S_{t-1}^D, S_t)P(S_t | S_{t-1}, S_{t-1}^D) \quad (6)$$

其中含 S_t^D 的分布 $P(S_t | S_{t-1}, S_{t-1}^D)$ 和 $P(S_t^D | S_{t-1}^D, S_t)$ 定义如下：

$$P(S_t = j | S_{t-1} = i, S_{t-1}^D = d) = \begin{cases} \delta(i, j), & \text{if } d > 0 \\ A(i, j), & \text{if } d = 0 \end{cases} \quad (7)$$

$$P(S_t^D = d' | S_{t-1}^D = d, S_t = k) = \begin{cases} p_k(d'), & \text{if } d = 0 \\ \delta(d', d-1), & \text{if } d > 0 \end{cases} \quad (8)$$

很显然如果 $P_i(d)$ 选择指数分布，则就等价于经典的 HMM 了。在加入显式段长变量后，就可以自由地选择 $P_i(d)$ 的分布，如采用参数形式如 Gaussian 分布、Gamma 分布等，从而只需要作较少的改动，就可以利用 DBN 通用的训练和识别算法。

4 实验与分析

实验是在连接数字语音库 Aurora-5^[9]上进行的，其内容是英文数字串，长度从 1 到 8 个不等，包括 zero 到 nine 以及 oh 这 11 个单词。选取 4 个说话人的语音样本，其中每人 125 句作为训练样本，15 句作为测试样本。由于是小词汇量语音识别，本文采用了整词模型，每个词由 8 个状态组成。另外，一个静音模型由 3 个输出状态组成，用来描述静音段。一个短停模型由 1 个绑定在静音模型中间状态的输出状态构成，用来描述语音信号中的换气等短停部分。采用 12 维的 MFCC 特征及其一阶和二阶差分，1 维的能量特征及其一阶和二阶差分，形成共 39 维的特征向量。统一使用 GMTK 软件包对所有模型进行训练和测试。

本文构建的基于段长分布的 DBN 语音识别模型如图 4 所示。图中加入了 Duration 这个显式的段长信息来控制状态的转移。为了易于一般化，同时引入了只取 0 和 1 值的二元变量 State transition 用来标识 Duration 是否递减为 0，即当 State transition 取 0 时，状态可以发生迁移，从而选择一个新的段长。

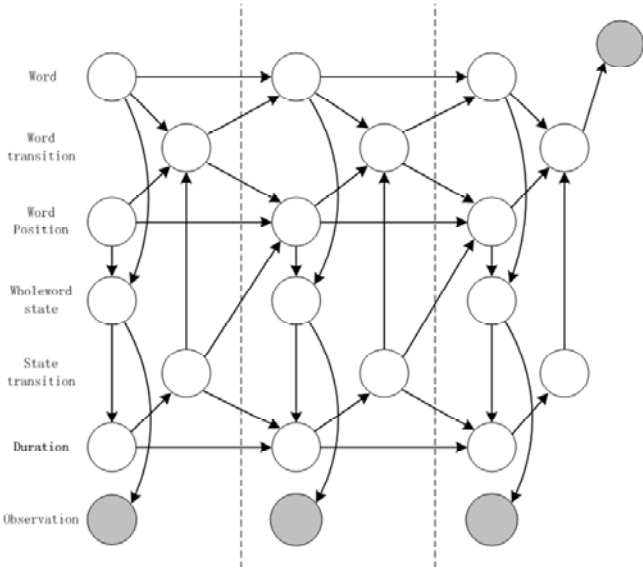


图 4 基于段长分布的 DBN 语音识别

为了验证段长对噪音的健壮性，对样本施加了不同信噪比的地铁噪音，实验结果如表 1 所示。

表 1 噪音环境下词识别错误率(WER)比较，其中 S, D, I 分别表示替换错误，删除错误，插入错误的个数

Model	clean				20 dB				15 dB				10 dB			
	S	D	I	WER	S	D	I	WER	S	D	I	WER	S	D	I	WER
HMM	5	0	21	10.70%	12	1	27	16.46%	13	2	40	22.63%	14	2	46	25.51%
DBN+duration	4	0	18	9.05%	9	1	28	15.64%	10	2	33	18.52%	12	3	36	20.99%

由表可知,从总体的词识别错误率(Word Error Rate, WER)看,含段长分布的 DBN 模型要低于 HMM 模型,尤其在噪音较大的情况下比较明显,这与段长对噪音健壮是吻合的。进一步观察替换错误(Substitution)、删除错误(Deletion)和插入错误(Insertion)三种错误类型可知,词识别错误率的改进主要归于插入错误的显著下降。在 15dB 和 10dB 噪音环境下,插入错误率分别相对下降了 17.5%和 21.74%。这是由于在 HMM 中段长呈指数下降,从而模型状态变迁较快,最终导致了较多的插入错误。

5 结论

本文针对 HMM 隐式段长分布的缺陷,研究了如何将段长显式地表示在 DBN 模型中。在 Aurora-5 语音库上的实验结果显示,基于段长分布的 DBN 比经典的 HMM 模型的识别效果有所改进,尤其在噪音环境下。

参 考 文 献

- [1] Vergyri D, Stolcke A, Gadde V R R, et al. Prosodic knowledge sources for automatic speech recognition[C]//Proc. of IEEE International Conference on Acoustics, Speech and Signal Processing. Hong Kong, China, 2003, 1:208-211.
- [2] Rabiner L R. A tutorial on hidden Markov models and selected applications in speech recognition[J]. Proc.of the IEEE, 1989, 77(2):257-286.
- [3] Ferguson J D. Variable duration models for speech[C]//Proc. of the Symposium on the Application of Hidden Markov Models to Text and Speech. Princeton, NJ, USA, 1980:143-179.
- [4] Sin B, Kim J H. Nonstationary hidden Markov model[J]. Signal Processing, 1995, 46(1):31-46.
- [5] Dean T, Kanazawa K. Probabilistic temporal reasoning[C]//Proc. Of The Seventh National Conference On Artificial Intelligence. Menlo Park, CA, USA, 1988:524-528.
- [6] Zweig G. Speech Recognition with Dynamic Bayesian Networks[D]. Berkeley, CA: Univ Californi California, Berkeley, 1998.
- [7] Bilmes J, Zweig G. The Graphical Models Toolkit: An open source software system for speech and time-series processing[C]//Proc. of IEEE International Conference on Acoustics, Speech and Signal Processing. Orlando, FL, USA, 2002, 4:3916--3919.
- [8] Friedman N, Murphy K, Russell S. Learning the structure of dynamic probabilistic networks[C]//Proc. Fourteenth Conference on Uncertainty in Artificial Intelligence. Madison, WI, USA, 1998:139-147.
- [9] Aurora-5[DB/OL]. <http://aurora.hsnr.de/aurora-5.html>

作者简介:

杜树木(1982-), 男, 江苏盐城人, 硕士生, 主要研究领域为语音识别;

何良华(1977-), 男, 讲师, 博士, 主要研究领域为模式识别、数字信号处理。

基于内容分析和点击率记录的混合音乐推荐系统

张 燕^{1,2} 李燕萍¹

(1. 南京理工大学计算机科学与技术学院, 江苏南京 210049;

2. 金陵科技学院信息技术学院, 江苏南京 211169)

摘 要: 随着网络技术不断发展, 人们希望以更快捷的方式获取喜欢的音乐, 功能强大的音乐推荐系统呼之欲出。目前已有的推荐方法主要依赖于点击率记录, 往往局限于与样本音乐相关的少数艺术家的作品, 这样的方式可以保证较高的命中率, 但是推荐范围太小。此外由于新作品的点击率很低, 系统基本不会推荐的新作品。而单纯基于内容的推荐系统由于特征描述的不稳定性会带来推荐成功率的降低。本文提出的基于内容分析和点击率记录的混合音频推荐系统, 通过对音乐内容特征的提取和描述, 结合传统推荐技术, 可以在更大范围内进行音乐片段的搜索, 降低作者、表演者及关联投票的限制。利用实验室采集的主观得分和主流音乐网站点击率数据进行实验, 结果表明, 混合推荐系统在推出新音乐的数量和推荐艺术家的多样性上大大超过了传统的方法, 并且保持了较高的命中率。

关键词: 音频推荐系统; 基于内容的音乐描述; 点击率纪录;

Hybrid Music Recommender System Based on Music Analysis and Rating Scores

Zhang Yan^{1,2} Tang Zhen min¹ Li Yan ping¹

1. Department of Computer S&T, Nanjing University of Science and Technology, Nanjing, Jiangsu, 210094, China; 2. Information Technology School of Jinling Institute of Technology, Nanjing, Jiangsu, 211169, China

Abstract: Along with the fast development of online service of large music collections, music recommender system is increasingly important to provide satisfactory access of music. However, traditional recommend method largely dependent on the same author, singer or other peoples rating scores, which will leader to narrow variety of artists and new music will be ignored. In this paper, we will propose a new hybrid music recommender system, which combines traditional method and the analysis of music content, to attain a higher recommendation accurate, with a wider variety of artist. We evaluate our system based on the rating scores and music from on-line music-service site. The result shows our new system will recommend music pieces from different artist and often obtain new productions, while with a high accurate.

Keywords: Music Recommender System; Content-based Music Description; Rating Scores;

1 引言

随着互联网技术的不断发展,海量的多媒体数据展现在用户面前,作为多媒体数据的一种,音乐成为人们生活中不可或缺的部分,然而面对互联网上数以万亿计的音乐数据,快速方便地获取喜爱的音乐成为用户迫切的需求,也成为音乐服务网站急需解决的问题。目前的解决方案主要包括检索和推荐两类。传统的检索方式主要通过用户提交音乐的标题、作者、表演者、专辑名或者歌词等文本信息进行查询^[1, 2],这需要用户对歌曲的标注信息非常熟悉。目前比较热门的新查询方式是基于内容的音乐检索,即通过描述音乐的风格、节奏、旋律等特征进行查询^[3, 4],这种方式需要用户具有一定的专业知识,并且没有建立通用的音乐内容描述规范,尚无法实现普及。一些先进的音乐服务站点提出了音乐推荐系统:利用用户已经和正在关注的音乐为样本,推荐类似音乐,并按优先顺序提供给用户选择试听,这种方式对用户没有专业要求,如果能保证推荐成功率,就会受到广大用户的欢迎。

然而,目前的音乐推荐系统主要是通过查询当前音乐同作者、同表演者、大风格相符的乐曲,或者提交喜欢当前音乐的其他用户的选择。例如当前用户喜欢歌曲 A,其演唱者 m 还表演过歌曲 C,而数据库中选择 A 的很多用户也选择了音乐 B,那么就将 B 和 C 推荐给当前用户。这两种方式,一般都能达到较高的命中率(即用户喜欢所推荐的音乐),然而这两种方法具有较强的局限性:前者基本上约束在少数几个艺术家上,后者对于新作品没有推荐能力(尚未获得足够点击率)。

基于内容的推荐系统不需要用户具有相关知识,通过主动分析样本音乐的特征,在较大艺术家范围内搜索推荐音乐,可以避免上述两种传统方法的不足之处^[5, 6]。但是基于内容的特征描述带有很多不确定性,推荐的命中率并不理想。此外,大量音乐数据的特征需要提取和存储,对时间和空间的需求都非常高。

为了解决上述问题,本文提出了同时包含上述方法的混合音乐推荐系统,其目的是在保证命中率的前提下扩大推荐音乐的搜索范围。本文希望利用神经网络的非线性能力将不同的描述方式统一在相同的模型中。实验结果表明该方法基本达到了上述研究目标。

2 音乐推荐算法

如前文所述,传统的推荐算法可以分为协同过滤算法和基于内容分析的算法。这些方法又可以进一步分为基于数据库的匹配方法和基于模型的分器方法。一般来说,分类器模型在训练时会消耗大量时间,一旦训练完成则可以很快实现任务目标;而匹配搜索方法则在推荐过程中需要较多的运算时间。

2.1 协同过滤法

1) 基于数据库的方法:该方法通过其他用户的访问记录预测当前用户对未知目标的兴趣程度。给定用户 u ,对目标音乐 m 的兴趣(用访问频率表示) $\tilde{r}_{u,m}$ 可以用下式表示:

$$\tilde{r}_{u,m} = \bar{r}_u + k \sum_{\{u'|u' \neq u, u' \in U\}} \omega_{u,u'} (r_{u',m} - \bar{r}_{u'})$$

其中, \bar{r}_u 和 $\bar{r}_{u'}$ 分别表示用户 u 和 u' 的平均访问频率, $\omega_{u,u'}$ 则用来描述用户 u 和 u' 在兴趣上的相似性, 最常用的定义为 Pearson 相关系数:

$$\omega_{u,u'} = \frac{\sum_m (r_{u,m} - \bar{r}_u) \sum_m (r_{u',m} - \bar{r}_{u'})}{\sqrt{\sum_m (r_{u,m} - \bar{r}_u)^2 \sum_m (r_{u',m} - \bar{r}_{u'})^2}}$$

利用该方式就可以预测当前用户未知音乐的兴趣程度, 从而实现音乐的推荐。

2) 模型法: 模型法主要从概率角度出发, 通过计算其他用户的期望得到音乐片段之间的相关性, 并通过当前用户所选择的若干音乐片段, 得到当前用户对音乐片段的兴趣分布, 从而得出相应的推荐曲目。预测过程可以表示为:

$$\tilde{r}_{u,m} = E(r|u, m) = \sum_r r \times p(r|u, m)$$

其中, r 表示主观得分 (0 到 4, 依次表示非常不喜欢、不太喜欢、还行、挺好的、非常喜欢), $p(r|u, m)$ 表示对用户 u , 观测音乐片段 m 时的兴趣程度。从这个角度来看, 对于特定用户, 具有一定风格的音乐会具有比较类似的主观得分, 因此训练出的模型就会对音乐类型相对敏感。所以这种方法推荐的多数是相同音乐风格的作品。

2.2 基于内容的音乐推荐方法

1) 基于数据库的方法: 通过查询与当前音乐内容相近的音乐片段推荐给使用者。其主要技术包括基于内容的音频特征描述, 特征空间的搜索等。给定音乐 m 的特征向量 Cm , M_u^+ 、 M_u^- 分别表示用户 u 喜欢和不喜欢的音乐特征集合, 则推荐算法如下:

第一步: 如果 M_u^+ 不为空, 则 M_u^+ 表示用户喜爱的音乐集合; M_u^- 表示不喜欢的音乐集合;

第二步: 分别计算每个候选项目的音乐特征 C 到目标集合 M_u^+ 和 M_u^- 的距离 $s_{u,m}^+$ 和 $s_{u,m}^-$ 。

第三步: 计算 $s_{u,m} = s_{u,m}^+ - s_{u,m}^-$ 作为最终得分, 推荐高分音乐片段。

2) 基于模型的方法: 音乐推荐可以看作将音乐片段分为喜欢和不喜欢两类的分类工作。因此可以引入分类器进行实现。假设基于内容的音频特征描述为矩阵 C , 则这个分类问题可以表示为:

$$p_+(m|u) = \prod_t p_+(t|u)^{C_{m,t}}$$

$$p_-(m|u) = \prod_t p_-(t|u)^{C_{m,t}}$$

其中, $p_+(m|u)$ 表示用户 u 对音乐片段 m 的兴趣, $p_-(t|u)$ 表示用户 u 对音乐特征 t 的兴趣。同样, 可以用下式得到最终得分, 并向用户推荐高分的音乐片段:

$$p(m|u) = \frac{p_+(m|u)}{p_-(m|u)}$$

2.3 混和音乐推荐算法

从上述方法中可以发现, 基于内容和基于记录的模型法都可以看作用户 u 和音乐片段 m

的概率函数，并通过加权计算或归一化处理得到最终的主观得分。其不同之处在于基于记录的方法依赖于其他用户的访问经历，而这种访问是有局限性的，因此搜索范围较小；而基于内容的方法依赖于音乐旋律等特征的提取，虽然搜索范围覆盖了数据库中所有的音乐片段，但旋律特征的表达具有不确定性，会影响最终的命中率。自然地，我们将二者结合起来，实现快速全面准确地数据库搜索和音乐推荐。最简单的方式是将二者加权相加，如下：

$$\alpha \times \tilde{r}_{u,m} + \beta \times p(m|u) \quad \alpha + \beta = 1$$

但在实际应用中，两个加权参数的选择很难确定，并且线性加权的方法掩盖了概率生成过程中的一些信息。因此我们考虑将 5 个基于记录的得分概率 $p(r|u,m)$ （用户 u 观测音乐片段 m 时的兴趣程度为 r 的概率），及基于内容的后验分类概率 $p_+(m|u)$ 、 $p_-(m|u)$ 共同作为输入形成一个神经网络识别器。利用神经网络的非线性能力和泛化性能实现系统输出的优化。结构图如下所示。

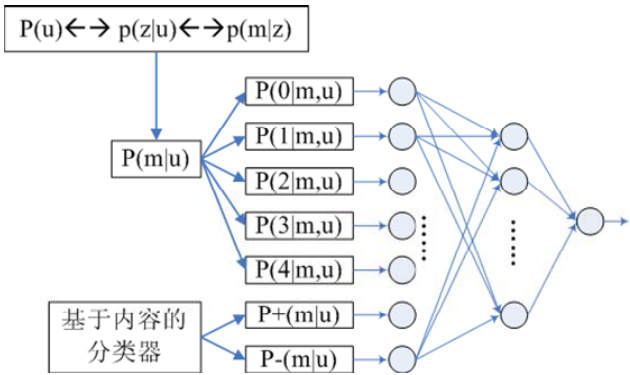


图 1 混合音乐推荐方法示意图

为了最终得到可排序的分值，神经网络的输出端采取了回归方式处理。整个网络将应用已有用户的选择和得分进行训练。而前期的输入数据获取则与传统算法法完全相同。由于训练好的神经网络在实际分类过程中只进行有限的加和乘的运算，因此附加的部分对原系统的处理速度影响很小。

2.4 基于内容的音乐特征提取

最后介绍基于内容的音乐特征的提取和描述。目前常用的基于短时帧的音频特征主要包括以下几类^[7]：

- 1) 频域能量： $E = \int_0^{\bar{w}} |F(w)|^2 dw$, $\bar{w} = fs/2$ ，若帧信号的能量小于给定阈值，则判定为静音帧。
- 2) 子带能量： $E = \int_{w \in sb_i} |F(w)|^2 dw$ ，将频谱分为几个子带计算能量可以表达能量在频谱上的分布。
- 3) 频率中心： $FC = \int_0^{\bar{w}} w |F(w)|^2 dw / E$ ，该指标可以描述音乐的亮度。
- 4) 带宽： $BW = \sqrt{\int_0^{\bar{w}} (w - FC)^2 |F(w)|^2 dw / E}$ ，用于衡量音频频域范围。

5) 基音频率：对于单独乐器（不含打击乐器）或人声，存在最小的谐波频率，可以表达音调高低。

6) MFCC 系数^[8]：根据 Mel 频谱变换计算信号的倒谱，符合人耳感知特性，主要表达声道特征或乐器的音色特征。

基于上述帧特征可以推导出基于音频片断的特征主要包括：

- ① 静音比例=静音帧数量/片段总帧数。
- ② 平均子带能量比：片段中各自带能量与频域总能量的比值的平均值。
- ③ 带宽均值：片断中各帧带宽均值。
- ④ 基音频率标准方差。
- ⑤ 谐成分比例：基频不为 0 的帧所占的比例。
- ⑥ 平滑基音比：基音平滑帧在音频片断中的比例。基音变化小于给定阈值的相邻帧称为基音平滑帧。

⑦ Δ MFCC：相邻帧 MFCC 的变化的均值，主要反映 MFCC 的动态变化特征，即音色的动态变化特性。

一般取 Δ MFCC 前 12 维即可，若子带数量为 5，则综合上述七类特征构成 22 维特征作为基于内容的音频特征参与后续分析和搜索。

3 实验、结果与分析

3.1 数据库获取和使用

对于本文系统来说，最好的实验数据是大量的用户对所有音乐片段都标注主观得分（喜爱的程度）。但是在实际应用中不可能获得这样的数据。因此我们选择了实验室标注和网络点击率评估相结合的方式获得数据库。首先我们确定 500 首热门流行音乐，然后根据其在主流音乐网站上的点击率为每个音乐进行评分。此外，我们选择在校大学生 120 人（无音乐背景）根据自己的喜好将 500 首音乐分为五类，每类 100 首歌曲。

利用这些数据，我们将对本文混合算法与传统方法进行对比，主要实验和对应数据包括：

1) 基于数据库的协同评估方法：80 人部分主观得分作为训练数据，40 人得分作为测试数据。此时数据将取出一部分得分大于 3 的音乐片段作为输入，观察剩余音乐得分与实际主观得分的差异性。

2) 基于模型的系统评估方法：利用每个音乐的点击率作为先验分布，取 100 人主观得分训练模型。采用剩余 20 人主观得分作为测试。

3) 基于内容的协同评估：提取部分音乐特征与对应用户得分相匹配构成数据库中的项，并根据该数据库搜索新用户对于未知音乐的兴趣程度。

4) 基于内容的模型评估方法：用 120 人的主观得分训练分类模型，并利用其中抽取的若干样本炎症模型评估的效果。

5) 混合推荐算法：利用 100 人若干首音乐的主观得分为目标输出，采用 2)、4) 提取相应概率参数，作为输入训练神经网络模型，用剩余音乐得分和剩余 20 人主观得分评估系统效果。

3.2 实验数据和分析

本节所有工作都基于上述五个系统的对比。首先针对首位推荐成功率、前三位推荐成功率和前十位推荐成功率进行实验。需要说明的是，当推荐的音乐片段主观得分等于 3 或 4 分时即认为推荐成功；此外，前 n 位推荐成功率指 n 个音乐片段都得到 3 或 4 分的比率，因此这个值一般会随着 n 的增加而减少。实验结果如下表所示。

从表中结果可以看出，基于数据库的方法推荐精度要大于基于内容的方法，这是与已有文献的结果相符合的。而我们的混合系统在前三位和前十位的推荐能力上都较其他四种传统方法有了较大的提高，但是在首位推荐成功率上不如基于数据库的模型方法，这说明利用其他用户记录和相同音乐家、音乐风格的约束，可以得到较好的首位匹配效果，但是随着搜索空间的扩大，其推荐能力将受到影响；而本文的混合算法结合了基于内容的分析方法一定程度上弥补了上述缺陷。

表 1 系统推荐成功率（%）对比和分析

位次	本文	基于记录		基于内容	
x	混合法	过滤法	模型法	过滤法	模型法
1	91.7	75.2	93.1	72.1	82.6
3	83.3	70.7	76.1	71.6	68.5
10	78.6	65.9	73.5	68.8	69.3

第二组实验数据主要反映本文算法在推荐不同表演者作品的的能力。传统的方法，特别是基于记录的方法依赖于其他用户之前的选择，而从用户自身来说，往往对表演者具有局限性，因此搜索结果倾向于较少的几个表演者的作品。本文算法利用音频内容特征，实现了大范围表演者作品的搜索。表 2 记录了不同推荐系统中表演者数量的比例，即推荐位次中表演者的数量与总推荐音乐片段数量的比值。

表 2 系统推荐表演者比率（%）对比和分析

位次	本文	基于记录		基于内容	
x	混合法	过滤法	模型法	过滤法	模型法
1	100	100	100	100	100
3	78.4	66.5	56.2	79.7	68.5
10	88.2	55.6	65.4	82.8	84.1

从数据中可以看出首位推荐的表演者只有一位，均为 100%，但随着位次的增加，可以看出，本文混合法推荐的表演者比例均为较高值，且随着位次增加呈上升趋势。这说明本文方法能够在较大的表演者范围内搜索音乐片段，并且保持较高的正确率。而传统的基于内容的方法虽然搜索范围也很大，但由于音乐特征的表现不稳定，正确率也会随着推荐位次的增加而不断下降。

表 3 系统推荐新音乐的比率（%）对比和分析

位次	本文	基于记录		基于内容	
x	混合法	过滤法	模型法	过滤法	模型法
1	0.3	0	0	0	0
3	7.4	0	0	9.1	6.7
10	6.9	0	0	10.4	9.8

第三组实验数据主要反映本文算法推荐新作品的能力，新作品指几乎没有其他用户点击记录的音乐片段。传统的基于记录的方法依赖于其他用户之前的选择，而新推出的作品基本上不可能得到较高的点击率，因此搜索就会忽略这些新作品。本文算法利用音频内容特征，实现了大范围表演者作品的搜索，包括了对新作品的分析。表 3 记录了不同推荐系统中新作品数量的比例，即推荐位次中新作品的数量与总推荐音乐片段数量的比值。从表中可以看出，基于内容的方法确实可以有效地搜索到相关的新作品，本文算法继承了这个特点，保证了新作品的推出。

4 结论和展望

针对当前基于数据库的推荐算法在表演者搜索范围过小，而基于内容的音乐推荐算法准确性较低的缺点。将两种方式有机地结合在一起，利用基于内容的特征扩大表演者的搜索范围，并利用数据库中的其他用户的信息规范搜索结果。在保证推荐成功率的前提下，获得了较大的表演者搜索范围，并且具有了推荐新翼乐片段的能力。实验表明，该系统在上述方面比传统方法有了较大的提高。但由于主观得分数据并不充分，网络点击率也是通过观测主流音乐网站评估得到，尚需要和相关领域进行合作测试实用效果。

5 致谢

本课题的研究得到江苏省现代教育技术研究所课题《基于内容的音频信息资源库构建与检索技术研究》（2007-R-4704）的资助与支持。课题研究过程得到音频处理课题组前成员钱博的大力支持，特别是在音频特征提取和检索算法的应用和实验过程中给予了很多建议，在此向他表示衷心的感谢！

参 考 文 献

- [1] A. Uitdenbogerd and R. van Schyndel, A review of factors affecting music recommender success, Proc. Int. Conf. Music Inf. Retrieval(ISMIR), 2002: 204-208.
- [2] G. Adomavicius and A. Tuzhilin, Toward next generation of recommender systems: A survey of the state-of-the-art and possible extersions, IEEE Tans. Knowl. Data Eng., vol.17, no.6, 734-749. Jun. 2005.
- [3] R. Typke, F. Wiering, and R. Veltkamp, A survey of music information retrieval systems, Proc. Int. Conf. Music Inf. Retrieval (ISMIR), 2005: 153-160.
- [4] B. Logan, Music recommendation from song sets, Proc. Int. Conf. Music Inf. Retrieval (ISMIR), 2004: 425-428.

- [5] U. Shardanand and P. Maes, Social information filtering: Algorithms for automating word of mouth, Pro. ACM Conf. Computer-Supported Cooperative Work, 1994: 175-186.
- [6] G. Linden and B. Smith, Amazon.com recommendations: Item-to-item collaborative filtering, IEEE Internet Comput., vol.7, no.1, pp.76-80, Jan. 2003.
- [7] 卢坚,陈毅松,孙正兴,张福炎,语音/音乐自动分类中的特征分析,计算机辅助设计与图形学学报, vol.14(3): 233-237
- [8] Rivarol Vergin, Douglas O'Shaughnessay. Generalized mel-frequency cepstral coefficients for large-vocabulary speaker-independent continuous speech recognition[J]. IEEE Trans. on Speech and Audio Processing, 1999, 7(5): 525-530

作者简介:

张燕,女,副教授,博士研究生,1969年出生,研究方向为基于内容的音频信息处理。

李燕萍,女,博士,1983年出生,研究方向为语音识别、说话人识别、音频处理技术。

能量重心校正法估计MPSK信号载波频差

张冬玲 葛临东

(信息工程大学信息工程学院, 河南郑州, 450002)

摘要: 研究在噪声中高精度估计 MPSK 信号载波频差的问题。新算法首先利用非线性变换使 MPSK 信号转化为单频信号, 然后加窗作 FFT, 用窗能量重心校正频率的方法, 达到精确估计载波频差的目的, 仿真表明新算法所需信号观察长度短、精度高、操作简单。

关键词: 载波频差估计; MPSK; 窗函数; FFT; 能量重心校正

Algorithm of Carrier Frequency Offset Estimation for MPSK Signal Based on Energy Centrobaric Correcting

Zhang Dong ling Ge Lin dong

(The College of Information Engineering, The Information Engineering University, Zhengzhou, Henan 450002, China)

Abstract: Deals with the problems of high accuracy estimation of carrier frequency offset for PSK signal in noise. A new method ,based on nonlinear algorithm is proposed which converts MPSK signal to single sinusoid signal ,and estimates the frequency offset accurately by using Energy Centrobaric Correcting Method. Simulation results show the new algorithm is feasible for high accuracy estimation of carrier frequency offset by using fewer data and is easy to implement.

Keywords: carrier frequency offset estimation ; MPSK; window function ; FFT ; Energy Centrobaric Correcting Method

1 引言

MPSK ($M=2, 4, 8$) 信号经过特定的非线性变换会产生正弦信号^{[1][2][3]}, 该正弦信号的频率为载波频差的 M 倍。例如 BPSK 信号经过二次非线性处理可产生频率为载波频差 2 倍的正弦信号, QPSK 信号经过四次非线性处理可产生频率为载波频差 4 倍的正弦信号。即可以通过非线性变换把 MPSK 类信号的载波频差估计问题转换为正弦信号的频率估计问题。

估计正弦信号的频率是信号处理领域一个热门话题, 方法有很多。基于 FFT 的离散频谱分析是一种较常用的方法, 但存在着各种误差, 目前国内外提出了许多校正方法^[4], 利用窗能量重心校正法^[5]就是其中的一种。该方法利用离散对称窗函数的能量重心的位置来校正频率、幅度和相位, 精度高, 且操作简单。

本文考虑的就是 MPSK 数字调制信号的载波频差估计问题。思路是对 MPSK 信号作非线性变换, 对信号加窗作 FFT, 用能量重心校正法估计载波频差, 体现算法思想的流程图如图 1

所示。文中以 QPSK 为例给出不同的信噪比、不同的偏差、不同的长度的 FFT 所得到的频差估计精度。仿真实验结果显示频率估计方差随着 SNR 的增加而减小,在 15dB 附近达到了 10^{-7} ,比未校正的方法提高了 3 个数量级; $|\delta|$ 越小,校正后方差比未校正的下降得越快;信号观察长度越短,新算法的改善效果越明显。总之,新算法所需信号观察长度短、精度高、操作简单。



图 1 能量中心校正 PSK 载波频差的流程图

2 能量重心校正PSK载波频差的原理

MPSK 信号波形可以表示为:

$$s(t) = Ae^{j2\pi(m-1)/M} e^{j(2\pi\Delta f_c t + \phi_0)} + z(t) \quad (1)$$

式中 $m = 1, 2, \dots, M$, A 为信号幅度, Δf_c 为载波频差, $\theta_m = 2\pi(m-1)/M$ 是载波的 M 个可能相位, 用于传发送信息, ϕ_0 为信号相位, $z(t)$ 时均值为零, 方差为 σ_z^2 的高斯白噪声。以采样频率 f_s 对 $s(t)$ 进行等间隔采样得到长度为 N 的序列 $s(n)$, 式(1)可以写成:

$$s(n) = Ae^{j2\pi(m-1)/M} e^{j(2\pi\Delta f_c / f_s n + \phi_0)} + z(n) \quad (2)$$

式中, $n = 0, 1, \dots, N-1$

对 $s(n)$ 作非线性变换, 以 QPSK ($M=4$) 为例, 用四次方变换即可。

$$r(n) = s^4(n) = A^4 e^{j(2\pi 4\Delta f_c / f_s n + 4\phi_0)} + z'(n) \quad (3)$$

显然, $r(n)$ 中出现了载波频差 4 倍的正弦信号, 可以用 FFT 的方法作频谱分析来估计出 Δf_c , 设 $r(n)$ 中待估计 $f_0 = 4\Delta f_c$ 。

对 $r(n)$ 作 FFT 前加窗 $w(n)$, 即

$$r_w(n) = r(n) \cdot w(n) \quad (4)$$

设窗函数的归一化幅度谱为 $W(\omega)$, 其归一化功率谱函数为 $P(\omega)$,

$$P(\omega) = W^2(\omega) \quad (5)$$

$r_w(n)$ 的归一化功率谱函数为:

$$P_{rw}(\omega) = P(\omega - \omega_0) \quad (6)$$

式中, $\omega_0 = 2\pi f_0$ 。

令 $\omega = \frac{2\pi k}{N}$, 则

$$P_{rw}(k) = P(k - k_0) \quad (7)$$

k_0 为 $P_{rw}(k)$ 幅度最大值对应的离散频率索引值(index)。只有当信号频率 f_0 为 FFT 频率分辨率 ($\Delta f = f_s/N$) 的整数倍时, 估计得到的频率值是准确的, 当信号频率与 FFT 的离散频率不重合时, 由于 FFT 的栅栏效应, 信号的实际频率位于两条谱线之间, 可将信号频率表示为

$$f_0 = (k_0 + \delta)\Delta f \quad (8)$$

δ 表示信号频率与 k_0 的相对偏差, 取值范围为 $-0.5 \leq \delta \leq 0.5$ 。显然仅利用 FFT 的幅度最

大值估计频率无法估计出 δ ，因而估计精度不高，要提高估计精度，只有增加信号的采样点数，这在某些应用场合是不可能实现的。因此研究一种信号点数不多情况下的高精度估计载波频差算法是很有必要的。

根据文献[4]中的证明有，只要是对称窗，当 n 足够大时，有：

$$\sum_{i=-n}^n P_{rw}(k_0 + i)(k_0 - k'_0 + i) = 0 \quad (9)$$

如图 2 所示， $k'_0 = k_0 + \delta$ 。

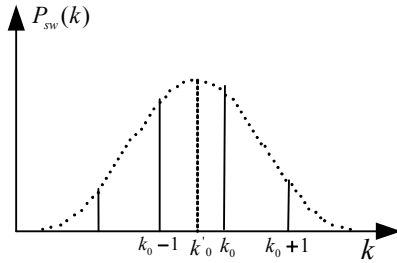


图 2 $s_w(n)$ 的功率谱

待估计频率 f_0 为：

$$f_0 = \frac{\sum_{i=-n}^n P_{rw}(k_0 + i)(k_0 + i)}{\sum_{i=-n}^n P_{rw}(k_0 + i)} \cdot \frac{f_s}{N} \quad (10)$$

设 $r_w(n)$ 的 FFT 为 $R_w(k)$ ，令 $n=1$ ， $A_{-1} = |R_w(k_0 - 1)|$ ， $A_0 = |R_w(k_0)|$ ， $A_1 = |R_w(k_0 + 1)|$ ，则式(10)可以写成：

$$f_0 = \frac{A_{-1}^2(k_0 - 1) + A_0^2 k_0 + A_1^2(k_0 + 1)}{A_{-1}^2 + A_0^2 + A_1^2} \cdot \frac{f_s}{N} \quad (11)$$

$$\delta = k_0 - k'_0 = \frac{A_1^2 - A_{-1}^2}{A_{-1}^2 + A_0^2 + A_1^2} \quad (12)$$

显然，若 k_0 位于窗能量重心位置，又因为窗的频谱是对称的，即 $A_1 = A_{-1}$ ，此时 $\delta = 0$ 。这与频谱分析的理论是吻合的。

3 仿真结果分析

Monte Carlo 仿真围绕不同的信噪比、不同的偏差 δ 、不同的信号长度之间的关系展开。3、4、5 给出了仿真结果，且每一数据的仿真次数为 1000 次。仿真数据由 MATLAB 产生，QPSK 的成形滤波器的滚降系数 $\alpha = 0.35$ ，每个符号有 4 个采样点。

第一组实验参数：设置 QPSK 信号的点数 $N = 1024$ ，采样频率 $f_s = 1024\text{Hz}$ ，载波频差 $\Delta f_c = 10.24\text{Hz}$ ，分别加 hanning 窗、hamming 窗和 blackman 窗，SNR 由 5dB~15dB 以步进 1dB 变化。得到的载波频差估计方差在不同信噪比下的性能曲线如图 3 所示。

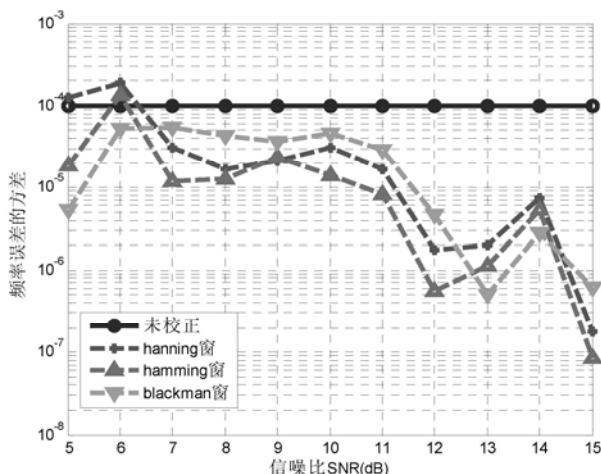


图3 随信噪比 SNR 载波频差估计方差变化规律

观察图3可以发现,随着 SNR 的变化,未校正的频率估计方差始终为 10^{-4} ,经过窗能量重心校正,频率估计方差随着 SNR 的增加而减小,在 15dB 附近达到了 10^{-7} ,比未校正的方法提高了3个数量级。

第二组实验参数:设置 QPSK 信号的点数 $N=1024$, 采样频率 $f_s=1024\text{Hz}$, 信噪比 $SNR=15\text{dB}$, 加 hanning 窗,载波频差估计的偏差 δ 在 -0.5 到 0.5 之间变化,步进为 0.1 。得到的载波频差估计方差在不同 FFT 点数比下的性能曲线如图4所示。

观察图4可以发现,随着 δ (信号频率与 FFT 幅度最大谱线所表示的频率之间的偏差)的变化,载波频差估计的方差不同, $|\delta|$ 越小,校正后方差比未校正的下降得越快,当 $\delta=0$ 时,频差落在 FFT 谱线上,此时未校正方法估计得误差为 0 ,所以图中出现断点。

第三组实验参数:设置 QPSK 信号的点数 $N=4096$, 采样频率 $f_s=1024\text{Hz}$, 载波频差 $\Delta f_c=10.24\text{Hz}$ 。对信号作 N 点 FFT,不做校正对频差进行估计。再对信号取前 1024 点用文中介绍的能量重心校正法进行频差估计(用 hanning 窗)。 SNR 由 $5\text{dB} \sim 15\text{dB}$ 以步进 1dB 变化。得到的载波频差估计方差在不同信噪比下的性能曲线如图5所示。

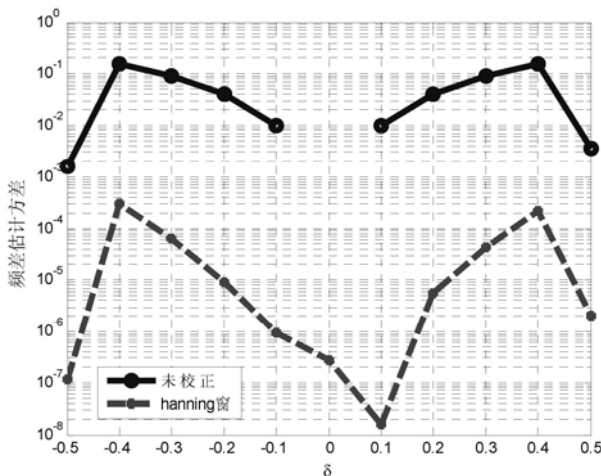


图4 随 δ 载波频差估计方差变化规律

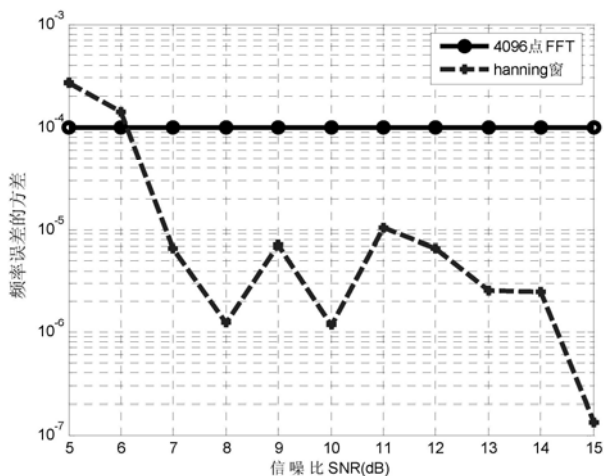


图 5 不同 FFT 点数的频率估计方差

观察图 5 发现, 随 SNR 增加, 能量重心校正法仅用 1024 点 FFT 得到的频率估计方差比 4096 点的 FFT 频率估计法的方差要小 2 个数量级以上。说明新算法适合用于信号观察长度较短的场合。

4 结论

本文直接利用 MPSK 类信号非线性变换的频谱结构特征, 给出了一种能量重心校正载波频差的算法。仿真结果表明新型算法能够很好地抑制噪声干扰, 而且信号观察长度越短, 新型算法的改善效果越明显, 所以该算法在高频谱利用率的突发通信载波频差估计方面具有明显优势。

参 考 文 献

- [1] Gardner W A. Exploitation of Spectral Redundancy in Cyclostationary Signals [J]. IEEE Signal Processing Magazine, 1991, 8(2), 14-19.
- [2] Reed D E and Wickert M A. Symbol-rate detection by a power-series-nonlinear envelope detector receiver[C]. Proceedings of the Seventh Annual International Phoenix Conference on Computers and Communications, 1988, Volume 1: 179-183.
- [3] Rappaport T S. Wireless communication: principles and practice [M]. Beijing: Publishing House of Electronics Industry, 1998, 229-233.
- [4] 丁康, 张晓飞. 频谱校正理论的发展. 振动工程学报, 2000.13(1):14 — 22.
- [5] 丁康, 江利旗. 离散频谱的能量重心校正法. 振动工程学报, 2001.9(3):354 — 358.

作者简介:

张冬玲 女, 1976 年 11 月生于江苏盐城, 讲师, 主要研究通信信号调制分析、软件无线电和盲源信号分离。

基于光流场法的目标跟踪算法在DM642 上的高效实现

武 曜 管 庆 徐 胜

(电子科技大学通信与信息工程学院, 四川 成都 610054)

摘 要: 本系统通过混合高斯背景建模产生背景, 然后检测出动目标, 再利用光流场法实现对物体的跟踪。在 DM642 上实现时, 使用了 TI 的 VLIB 视频处理库, 大大提高了代码的执行效率。利用 DM642 的硬件特点对程序进行优化, 进一步提高代码效率。最后通过实验证明, 在 DM642 上可以实现基于光流场法的多目标跟踪。

关键词: 物体跟踪; 光流场法; 混合高斯背景建模; VLIB; DM642

Implementation of the Object Tracking Algorithm Based On Optical Flow

Wu Yao Guan Qing Xu Sheng

(School of Communication and Information Engineering ,University of Electronic Science and Technology of China, Chengdu, 610054, China)

Abstract: Modeling through a multi-Gaussian background to detect moving objects. And then using optical flow method to track moving objects. The use of TI's VLIB image processing library, greatly improving the efficiency of the implementation code. In addition, the use the characteristics of DM642 hardware to optimize the code to further improve the code efficiency. The experiments show that the algorithm in the DM642 can archive real-time object tracking.

Keywords: Object tracking ; Multi-Gaussian Background Modeling; Optical flow; VLIB; DM642

1 引言

光流场^[1]是指图像灰度模式的表观运动。它是一个二维矢量场, 它包含的信息即是各像素点的瞬时运动速度矢量信息。研究光流场的目的就是为了从序列图像中近似计算不能直接得到的运动场。

但是由于光流场算法的复杂度大, 所以到目前为止在 DSP 上实现此算法的系统还很少。随着 DSP 处理器的发展, 以及很多高效的算法库的出现, 在 DSP 上实现使用光流场法对动目标进行跟踪成为了可能。

本文将详细阐述如何在 TI 公司的 DM642 上应用光流场方法高效实现多目标物体跟踪。

2 物体跟踪方案

2.1 系统组成

整个系统结构如图 1 所示。首先，输入图像将会先用来建立背景。其中，背景使用了混合高斯模型进行建模，然后使用背景差分的方法进行动目标提取，当前帧与背景帧像素之间的差值大于预先设定的阈值时，会将该像素点判定为动目标。接着使用标号算法对动目标进行标号，标号算法将图像中的每一个物体进行编号，并返回物体位置。最后对连续图像中的同一物体进行跟踪，实现动目标的跟踪。

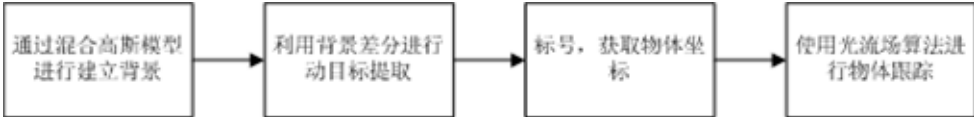


图 1 系统结构图

2.2 算法介绍

根据上一节所表述的系统结构，所对应的算法包括了背景建立算法，提取动目标算法，标号算法以及跟踪算法。下面将分别介绍这几种算法。

2.2.1 混合高斯背景建模

背景建立的方法有很多种，比如固定背景，多帧平均，单高斯背景建模，混合高斯背景建模等方法。在实现中，我们将使用具有自适应，拥有多模态的混合高斯背景建模^[1]方法。

1) 算法概述

混合高斯背景建模的基本思想是为每一个像素建立一个混合高斯模型。根据模型中高斯分布的权值以及方差选取可能的背景分布。如果像素点的值符合任意一个背景分布，则判定为背景点；否则，视为动目标。

混合高斯背景建模的步骤一般为：混合高斯模型的建立及更新，背景分布的选取，动目标像素的检测。

2) 混合高斯模型的建立及更新

我们可以将某一个点在一段时间的值视为一个“像素过程”。像素过程是一个像素值的时间序列。在任何一个时间 t 对于一个点 $\{x_0, y_0\}$ 所知道的是它的历史

$$\{X_1, \dots, X_t\} = \{I(x_0, y_0, i) : 1 \leq i \leq t\} \quad (1)$$

式(1)中的 I 为图像序列。

而对于每一个像素点 $\{X_1, \dots, X_t\}$ 则被建模成一个由 K 个高斯分布组成的混合高斯分布。当前的像素点概率值可能为

$$P(X_t) = \sum_{i=1}^K \omega_{i,t} \times \eta(X_t, \mu_{i,t}, \sum_{i,t}) \quad (2)$$

其中 $\omega_{i,t}$ 是一个在 t 时刻混合高斯模型中第 i 个高斯分布的权重。 $\mu_{i,t}$ 是 t 时刻第 i 个高斯

分布的期望， $\sum_{i,t}$ 是第 i 个高斯分布在时间 t 的协方差矩阵， η 是高斯概率密度函数

$$\eta(X_t, \mu, \sum) = \frac{1}{(2\pi)^{\frac{n}{2}} |\sum|^{\frac{1}{2}}} e^{-\frac{1}{2}(X_t - \mu)^T \sum^{-1}(X_t - \mu)} \quad (3)$$

对于获得的输入图像，将像素的值与 K 个高斯分布进行匹配。然后，根据匹配结果进行模型的参数更新^[2]，以保证算法对于环境变化的自适应性。

3) 背景分布的选取

针对更新后的模型，需要从 K 个高斯分布中选取可能的 B 个背景分布建立背景模型。具体步骤如下：

第一步，将 K 个高斯分布根据 ω_i / σ_i 的值按从大到小的顺序排列，其中 σ_i 为 t 时刻第 i 个高斯分布的方差；

第二步，选取 B ($B \leq K$) 个分布建立背景模型，其中

$$B = \arg \min_b (\sum_{k=1}^b \omega_k > T) \quad (4)$$

其中 T 是一个被认作是背景的一个最小值。

4) 动目标像素的检测

将当前像素的值与背景模型进行比对。如果像素属于背景模型的任意一个分布，则判断为背景点；否则判断为动目标点。

2.2.2 标号算法

经过动目标提取后，得到的动目标掩码通常包含多个连通区域。其中，目标掩码指动目标所对应的像素点为 1，背景所对应的像素点为 0 的图像。标号算法根据预先设定的连通结构（四连通或八连通）对各个区域进行标号和自动识别，从而获得动目标的坐标和形状等特性。

2.2.3 光流场法

光流场法在模式识别，计算机视觉以及其他图像处理应用中非常有用。它与运动检测以及运动估计紧密相关。而在一个图像平面上，物体的运动往往是通过图像序列中不同点的灰度分布的变化体现的。从而，空间中的运动场转移到图像上就表示为光流场，光流场反映了图像上每一点的灰度的变化趋势。

2.2.4 Lucas-Kanade 光流场法

Lucas-Kanade 是光流场法中的一种，此方法仍然是运动估计中最流行的两帧差分算法。这种方法假设以一点为中心的区域中的光流是不变的。Lucas-Kanade 方法以光流场法为基础。预测所需的附加的限制条件为以 $m \times m$ ($m > 1$) 的区域中的光流 (V_x, V_y) 为一常量，根据此条件所列的条件如方程组(5)所示。

$$\begin{aligned}
I_{x1}V_x + I_{y1}V_y &= -I_{t1} \\
I_{x2}V_x + I_{y2}V_y &= -I_{t2} \\
&\vdots \\
I_{xn}V_x + I_{yn}V_y &= -I_{tn}
\end{aligned}
\tag{5}$$

解此方程的方法通常为最小均方根方法：

$$\begin{bmatrix} V_x \\ V_y \end{bmatrix} = \begin{bmatrix} \sum I_{x_i}^2 & \sum I_{x_i} I_{t_i} \\ \sum I_{x_i} I_{t_i} & \sum I_{y_i}^2 \end{bmatrix}^{-1} \begin{bmatrix} -\sum I_{x_i} I_{t_i} \\ -\sum I_{y_i} I_{t_i} \end{bmatrix}
\tag{6}$$

即可计算出相应像素点的光流值，根据光流值即可预测出在下一帧图像中此像素点出现的位置。

3 算法实现

3.1 程序流程

程序流程如图 2 所示。

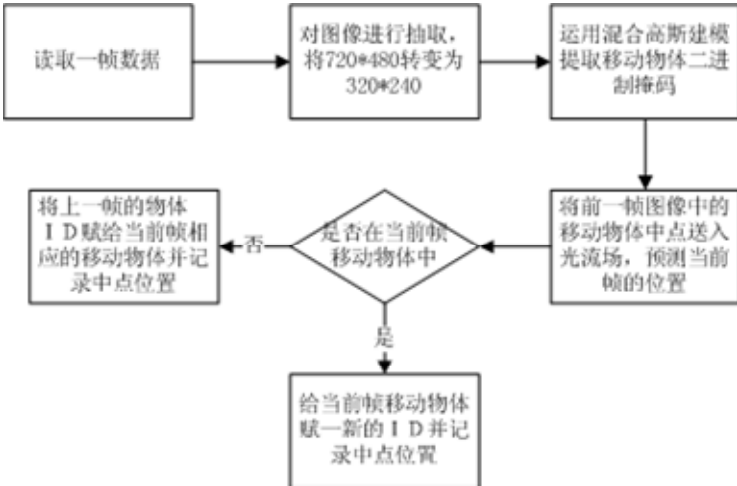


图 2 程序流程图

首先对输入图像使用图像金字塔函数，将原来的 640×480 的图像抽取成为 320×240 的图像，用以减小每帧图像的运算量；其次将抽取后图像的 Y 分量输入到高斯背景提取动目标掩码函数，得到当前帧的二进制掩码。然后将其输入到框取函数中得到当前帧的动目标个数以及具体位置数据。第四步将得到的动目标的中心坐标输入到光流场函数中，预测出下一帧图像中此中心点坐标的位置，如果预测出的物体位置正好处在下一帧某一动目标的范围之内，则判断此物体与前一帧为同一物体，并赋予此物体相同的 ID 号。如果没有搜索到相应的物体，则表示赋予其一个新的 ID 号，其中，物体的 ID 从 1 开始，每增加一个 ID，其号码为前一个物体 ID 号加 1。最后，在一个图像序列中具有相同 ID 号的物体表示同一物体，实现了动目标的跟踪。

3.1.1 物体数据结构体

为了跟踪物体，这里将建立一个存储物体相关数据的结构体。结构体中包括物体 ID，物体中心点坐标。其中，每个物体的 ID 号是唯一的一个整型数。每出现一个新的物体，其所对应的 ID 将为前一个物体的 ID 号加 1。变量 ID 的初始值为 0。

在程序最开始，先定义一个结构体数组，数组的大小将决定能够跟踪对象的个数，但是数组越大，程序执行速度也将越慢。

在主循环中，程序先对当前帧进行混合高斯背景建模，然后分析是否有动目标，如果当前帧中有动目标，则先对结构体数组中非零 ID 的中心点坐标进行光流场预测，预测出在当前帧中物体可能出现的位置。得到预测到的坐标点后，再判断该点是否属于当前帧的动目标，如果属于，则更新相应结构体元素的中点坐标。否则，在结构体数组中找到一个 ID 为 0 的元素，将其 ID 号赋值并记录物体中心点的值。

循环上述过程，完成物体的跟踪。

当连续 5 帧图像某 ID 物体没有出现时就认为此物体在图像中消失，将相应的结构体数组元素 ID 至为 0。

4 程序优化

代码的优化包括三个层次：算法的优化，程序本身的优化，利用 DSP 的结构对代码进行优化。

由于在实现中我们将使用 VLIB 所提供的算法，所以这里将只针对后两个方面进行优化。

4.1 VLIB

VLIB^[5]是由 TI 公司免费提供的用于视频处理的代码库，其中包含了许多主流的视频处理算法。

4.1.1 函数介绍

为了高效实现物体跟踪算法，我们利用了 VLIB 中提供的几个函数，其中包括混合高斯背景提取动目标掩码函数，光流场法函数，图像竖直方向以及水平方向梯度的计算函数，图像金字塔函数以及解压缩函数下面将详细介绍这几个函数的作用：

1) 高斯背景提取动目标掩码函数 VLIB_mixtureOfGaussiansS16：此函数用于对 YUV 图像和的 Y 分量进行分析，并提取出图像中的动目标。并返回一个经过压缩的二进制数组。其中，数据压缩是指，由于返回的动目标掩码只有两个数据 0 与 255，所以将一个使用 8 比特表示的像素点压缩成为只使用 1 个比特表示的数据。这样做使得运算效率大大提高。

2) 光流场法函数 VLIB_trackFeaturesLucasKanade_7x7：输入为当前帧、前一帧图像、当前帧图像的水平方向梯度以及竖直方向梯度、前一帧图像中坐标点的坐标。输出为当前帧中预测的属于前一帧输入坐标的坐标位置。

3) 图像竖直方向以及水平方向梯度的计算函数 VLIB_xyGradientsAndMagnitude：此函数输入为当前帧图像，输出为当前帧图像的水平方向梯度以及竖直方向梯度。

4) 图像金字塔函数 VLIB_imagePyramid8：此函数的作用是将一幅较大的图像抽取成为

一幅尺寸较小的图像，使用此函数的目的是减小 DSP 处理数据的量，这样可以使得处理图像的帧率提高，以使得系统对事件的响应时间缩短。

5) 解压缩函数 `VLIB_unpackMask32`: 正如前面所述，为了提高 DSP 运算效率，VLIB 算法库对每帧数据进行了压缩，使得原来使用 8 比特表示的像素点压缩成为 1bit 表示。但是最终的图像显示必须使用 8bit 数据进行表示，所以必须使用加压缩函数将经过压缩的图像数据进行解压，以进行显示等的后续应用。

6) 框取函数 `VLIB_createConnectedComponentsList`: 此函数输入为由高斯背景提取动目标掩码函数输出地二进制掩码数据，经过此函数的处理可以得到当前帧中的动目标的个数、位置以及外接框等数据。

4.2 缩小处理图像尺寸

由于图像的尺寸对于处理速度有很大影响，比如一个 640×480 的图像要比一幅 320×240 的图像多处理 3 倍数据量，所以减小图像尺寸将大大提高处理效率。

但是，缩小图像尺寸后也会带来副作用，即会丢失一些图像的细节信息。这就需要在准确度和处理速度上间权衡。在实现中，由于监控对象为人车等大物体，所以适当的缩小图像尺寸来换取速度是可行的。

在这里，我们将 640×480 的图像抽取成 320×240 的图像。由于在抽取时采用对一个 2×2 的区域选取一个像素点的方法实现，所以在处理完成后，物体坐标只需要乘以 2 即可恢复出在原图像中的坐标。

4.3 展开循环

由于每一次循环完成后都要进行跳转，这样是会影响程序执行速度的。一种解决方法就是将循环展开，也就是将一个循环中的代码直接再追加相同的代码而不使用循环。

但是在打开编译器的优化开关时，编译器并不会展开循环，因为循环次数为偶数与奇数的展开方法不同，当编译器不知道循环的循环次数时是不会展开的。所以程序员需要使用 `MUST_ITERATE` 指令来说明循环体的结构，让编译器帮助展开循环，以提高代码效率。其中 `MUST_ITERATE` 为告知编译器循环体的具体结构的指令。

4.4 对CACHE的优化

4.4.1 CACHE

由于 DSP 的内部存储器资源有限，程序代码及数据不可能全部放到 DSP 的片上存储器运行，所以它在片上设计了一块存储器配合片外慢速存储器，当 DSP 需要执行指令或数据的时候会自动将数据搬移到 `CACHE`^[5]中，由于 `CACHE` 属于片上存储器所以可以大大提高数据的处理速度。

4.4.2 CACHE MISS

当 DSP 执行程序或进行计算时，所需要的数据不在 `CACHE` 中，这时 `CACHE` 需要从外部存储器读入 DSP 所需要的数据，这时就产生了一个 `CACHE MISS`。

在图像处理中，经常遇到的一种 CACHE MISS 的情况是：如果图像以列方向来处理图像的话，在遇到一个 MISS 后 CACHE 会连后面的数据也读入，而图像以行方向来处理时由于数据是不连续的，每个像素数据之间相差图像的列个数，如果两个元素之间的地址相差大于 CACHE 一次读入的数据量 128 字节时，尽管当 CACHE MISS 后，也读入连续的地址的数据，但是并没有读入程序处理所需的数据，故一直产生 CACHE MISS，使得 CACHE 没有发挥任何作用，从而让图像处理效率低下。

4.4.3 CACHE具体优化方法

为了解决上一节中提到的问题，在程序中可以做如下处理。将一个循环展开成为几个循环。

如程序

```
for (i = 0; i < MAT_ORDER; i++) {
    for (j = 0; j < MAT_ORDER; j++){
        C[i][j] = mat_oprn(A[i][j],B[j][i]);
    }
}
```

修改为

```
/* Matrices are accessed in parts and the parts are reused */
for (ii = 0; ii < MAT_ORDER; ii += MAT_PART_SIZE) {
    for (jj = 0; jj < MAT_ORDER; jj += MAT_PART_SIZE) {
        for (i = ii; i < MAT_PART_SIZE + ii; i++) {
            for (j = jj; j < MAT_PART_SIZE + jj; j++){
                C[i][j] = mat_oprn(A[i][j],B[j][i]);
            }
        }
    }
}
```

将图像数据的访问模式修改为如上所示以后，图像数据的将会分成很多子块进行访问，这样做可以使得每个子块得到重复使用，以提高 CACHE 使用效率，减少 CACHE MISS 的发生。

5 实验结果

实验使用 DM642 作为硬件平台，并对一幅停车场的视频进行测试。输入系统的图像尺寸为 720×576，帧率 25 帧每秒。视频中最多出现两个动目标，其中包括人及车。此试验的目的为了跟踪视频中出现的的所有动目标,如果成功跟踪，那么被判定成相同动目标的外接框将用相同的颜色标明。实验结果如图 3 所示。



图 3

左图为第 1 帧跟踪到的两个物体，右边的图像为 30 帧以后跟踪到的物体，其中用相同颜色的框来框取同一物体。蓝框和红框分别稳定地跟踪了行人和白色车辆，实现多目标跟踪。

但是，在多次试验后发现，在跟踪过程中也偶尔出现跟踪失败的情况，即将连续两帧中的同一物体判断为两个不同的物体。原因是只取外接框的中心点来作为跟踪点的话，很有可能不在动目标上。

为了提高系统的处理帧率，对系统代码进行了相应的优化。主要方法包括缩小图像、利用 CCS 对循环的展开以及对于 CACHE 的优化。下表显示了在 DSP 使用效率为 95%时加入优化后的帧率提升（右侧的优化项目为在左侧优化项目的基础上进行的）

表 1 优化帧率统计表

优化项目	无优化	缩小图像	展开循环	优化 CACHE
帧率	1 帧/s	8 帧/s	10 帧/s	15 帧/s

从表 1 可以看出，经过较细致的优化之后，系统的处理帧率达到了 15 帧每秒，具有很好的实时性。

6 结论及未来的工作

文章介绍了一种多目标跟踪的实现方法，包括了混合高斯模型对背景建模，使用标号算法得到物体位置以及运用，用光流场法对多目标进行跟踪。我们也在 DM642 上高效的实现了该跟踪算法，从实验结果可以看出，本方法可以对多目标进行跟踪，但是跟踪准确率还有待提高。根据测试，物体一般在 100 帧左右会跟踪失败。最后，通过对实现代码的优化，达到了实时监控的要求。

未来的工作将集中在提高跟踪准确率上，可以考虑将包含物体更多信息的特征点（如 Harris 角点）作为跟踪点，也可以增加跟踪点的数目。

参 考 文 献

- [1] Lucas B D and Kanade T, “An iterative image registration technique with an application to stereo vision. “ in Proceedings of Imaging understanding workshop, pp 121—130, 1981
- [2] Chris Stauffer, W.E.L. Grimson, “Adaptive background mixture models for real time tracking,” in CVPR, 1999.
- [3] Ince, S., Konrad, J, “Occlusion-Aware Optical Flow Estimation,” in IEEE, 2008.
- [4] 彭启琮,管庆. DSP 集成开发环境—CCS 及 DSP/BIOS 的原理与应用. [M]北京:电子工业出版社, 2005 年
- [5] TMS320C64x+ DSP Cache User’s Guide, SPRU862A, October 2006.
- [6] TMS320C6000 Optimizing Compiler v6.0 Beta, SPRU187N, July 2005.

作者简介:

武曜 (1983-), 男, 山西太原人, 硕士, 主要研究领域为信号处理;

管庆 (1968-), 硕士, 男,副教授, 主要研究领域为信号处理、DSP 技术应用;

徐胜 (1978-), 男,博士, 主要研究领域为图像识别。

目标跟踪算法在DM6437 上的实现和优化

赵 渊 管 庆 徐 胜

(电子科技大学通信与信息工程学院, 四川 成都 610054)

摘 要: 在高性能的 DM6437 上, 利用 TI 公司的 VLIB 库和 IQmath 库, 实现了结合 Mean Shift 与卡尔曼滤波的目标跟踪算法。通过混合高斯建模、腐蚀、膨胀、目标检测, 结合 Mean Shift 和卡尔曼滤波算法对动目标进行跟踪。经过多种不同层次的优化策略, 在 DM6437 上高效地实现了该跟踪算法。实验结果表明, 该实现具有良好的性能, 并且满足高效实时性。

关键词: 算法实现与优化; Mean Shift; 卡尔曼滤波; DM6437;

Realization and Optimization of the Object Tracking Algorithm on DM6437

Yuan Zhao Qing Guan Sheng Xu

(School of Communication and Information Engineering, UESTC, Chengdu, 610054, China)

Abstract: Using TI's VLIB and IQmath libraries, an object tracking algorithm combined Mean Shift with Kalman filtering is implemented on the high-performance DM6437. Through Gaussian Mixture Modeling, erosion, dilation and moving object detection, the tracking object is automatically selected and tracked by combining Mean Shift with Kalman filtering algorithm. With several kinds of different levels of optimization strategy, this tracking algorithm has been efficiently realized on DM6437. The experimental results show that the achievement of good performance, and meet the high-performance real-time.

Keywords: Algorithm Realization and Optimization; Mean Shift; Kalman Filter; DM6437;

1 引言

1995 年 Cheng^[1]将 Mean Shift 引入计算机视觉领域后, 在跟踪方面得到了广泛应用。它采用核函数直方图建模, 对边缘阻挡、目标的旋转、变形以及背景运动都不敏感, 而且它的算法实时性好, 有利于在嵌入式处理器上的实现。但当目标运动速度过快以及发生遮挡等问题时, 其跟踪效果不好, 需使用卡尔曼滤波器进行目标位置的预测。而它们需要处理的视频数据量大、运算复杂, 因此需要高性能的 DSP 才能进行实时处理。于是本文选用 TI 的高性能视频处理芯片 DM6437 实现对视频目标的实时跟踪。

2 算法介绍

2.1 算法流程简介

本文进行算法实现时对视频处理的流程如图 1 所示，首先对采集的视频图像采用混合高斯模型算法进行背景建模，然后对二值图像进行形态学滤波，接下来是动目标检测和选取待跟踪目标，最后结合 Mean Shift 与卡尔曼滤波对目标进行实时跟踪。

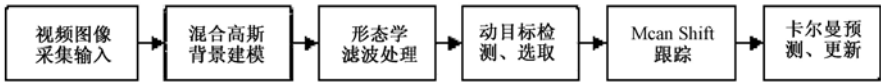


图 1 算法流程图

2.2 混合高斯背景建模

背景的复杂性往往会给动目标检测带来很大困难，比如光线变化，树叶摇曳，阴影干扰等，这些都影响到跟踪的准确性和可靠性。在这么复杂的环境中需要设计一种好的背景模型和背景差分算法来抽取运动区域，于是本文采用混合高斯背景建模算法^[2]，该算法能很好地区别导致每一帧视频图像中像素点变化的原因是背景物体的变化还是前景物体的运动。

2.3 形态学滤波

经过混合高斯背景建模，得到的二值图像中的前景区域不一定是真正的动目标。由于噪声的干扰以及目标与背景图像之间往往有小部分颜色和灰度相似，那么二值图像中通常含有许多孤立的点、孤立的区域、零星的空穴等等。因此一方面需要将属于运动目标的离散点连接起来，另一方面又要去除噪声点。对二值图像进行膨胀和腐蚀形态学滤波来消除这些干扰，去除存在的小噪声块，同时填补运动目标内可能因为漏检而出现的空穴。

2.4 动目标检测

前景动目标大多数表示为连通组件或者图像像素块。检测算法能产生一个二进制掩码，该掩码能够识别每个像素点到底是属于前景还是背景。本文采用的连通组件算法能够把离散的区域标记为多个目标块，通过对目标块的分析，能够提取出有用的目标信息，比如目标的质心、目标矩形框和 ID 号，作为目标跟踪所需要的特征。

2.5 Mean Shift与卡尔曼滤波相结合算法

经过动目标的检测，得到了动目标的块状区域，接下来就是做进一步处理——跟踪。本文采用 Mean Shift 和卡尔曼滤波器^[3]相结合的算法，下面详细介绍该算法的具体过程。

2.5.1 目标模型和候选模型建模

本文采用动目标检测自动选取动目标。在检测出的运动区域内选取一个面积较大的作为

目标区域，将其从当前图像中分割，作为目标模型。然后建立候选模型，对动目标在接下来每帧中可能包含的区域的描述称为候选模型。目标模型和候选模型的建模如下：

$$\hat{q}_u = C \sum_{i=1}^n k(\|\frac{x_0 - x_i}{h}\|^2) \delta[b(x_i) - u] \quad (1)$$

$$C = 1 / \sum_{i=1}^n k(\|\frac{x_0 - x_i}{h}\|^2) \quad (2)$$

$$\hat{p}_u(y) = C_h \sum_{i=1}^{n_h} k(\|\frac{y - x_i}{h}\|^2) \delta[b(x_i) - u] \quad (3)$$

$$C_h = 1 / \sum_{i=1}^{n_h} k(\|\frac{y - x_i}{h}\|^2) \quad (4)$$

这里， \hat{q}_u 表示目标模型， $\hat{p}_u(y)$ 表示在候选点 y 处的候选模型。 x_0 表示目标区域的中心，目标区域中有 n 个像素用 $\{x_i\}_{i=1 \dots n}$ 表示，特征值 bin 的个数为 m 个。 $k(x)$ 为核函数的轮廓函数，本文选择 Epanechnikov 函数。由于遮挡或者背景的影响，目标模型中心附近的像素比边缘的像素更可靠， $k(x)$ 对中心的像素给一个大的权值，而远离中心的像素一个小的权值。函数 $k(x)$ 中 $\|\frac{x_0 - x_i}{h}\|^2$ 的作用是为消除不同大小目标计算时的影响。 $\delta[b(x_i) - u]$ 的作用是判断目标区域中像素 x_i 的颜色值是否属于第 u 个 bin，属于则为 1，否则为 0。

2.5.2 相似性函数计算

本文使用巴氏系数作为描述目标模型与候选模型之间相似度的相似性函数，其定义为：

$$\hat{\rho}(y) = \sum_{u=1}^m \sqrt{\hat{p}_u(y) \hat{q}_u} \quad (5)$$

相似性函数值在 0~1 之间，该值越大，表示两个模型越相似，在当前帧中不同的候选区域计算得到的候选模型，使它最大的候选区域即是目标在本帧中的最终位置。

2.5.3 目标定位

为使相似性函数值最大，在当前帧的目标中心先定位为前一帧中目标中心的位置 y_0 ，从这一点寻找局部最优目标位置 y_1 。通过对相似性函数求最大值，可得到候选区域中心 y_0 移动到真实目标区域 y 的 Mean Shift 向量：

$$m_{hG}(y) = y_1 - y_0 \quad (6)$$

Mean Shift 算法反复迭代，最后得到在当前帧目标的最终位置 y 。

2.5.4 卡尔曼滤波器预测Mean Shift起始点

由于相似性函数在 y_0 处进行泰勒展开，这就限制了起始点 y_0 和 y 点的距离不能太大。如果目标运动过快，Mean Shift 跟踪效果不好。因此采用卡尔曼滤波器来解决该问题。利用前 $k-1$ 帧中目标中心的位置信息作为卡尔曼滤波器的观测值，首先用卡尔曼滤波器预测 k 帧目标的位置，作为 Mean Shift 算法中目标的起始点，Mean Shift 会在这点的邻域内找到目标最优的位置，再以该点作为卡尔曼滤波器的观测值，进行下一帧的运算。

3 算法实现

本文使用 TI 公司的 DM6437EVM 和 DVSDK 软硬件平台进行算法的实现和优化。DM6437^[4]的 DSP 核上运行 DSP/BIOS 实时操作系统^[5]、背景建模算法、目标检测算法、跟踪算法。程序从 Main（）开始，系统依次完成各个软硬件模块的初始化，建立视频输入、视频编码输出和 OSD 输出通道，并且获取第一帧图像。完成这一系列工作以后，Main（）返回。系统进入 DSP/BIOS 调度时间，让视频处理主线程来完成一帧一帧的视频分析，最终将跟踪结果通过视频输出通道送到屏幕上。

视频处理主线程是整个程序的核心，主要包括混合高斯建模、形态学处理、动目标检测、Mean Shift 跟踪和卡尔曼预测与更新五个模块。由于本文关注的是视频目标的实时跟踪，因此关于混合高斯建模、形态学处理、动目标检测等算法，在实现的时候本文采用了 TI 的 VLIB 库中 API 函数来完成。这些 API 函数都是经过高效优化的，为后续的跟踪算法的处理节省了大量的时间和资源，使得整个实现能够获得高效和实时性。

下面详述 Mean Shift 和卡尔曼滤波相结合算法的实现步骤，其流程设计如图 2 所示。

- 1) 自动初始化：利用动目标检测算法得到目标块信息（包括目标的心形坐标和目标块的大小），计算目标模型的 RGB 颜色直方图；
- 2) 在下一帧视频图像中，利用卡尔曼滤波器的状态向量中的 x 与 y 方向的位置和速度信息来预测当前帧中候选模型的起始点；
- 3) Mean Shift 迭代：利用三种不同大小的目标搜索窗在相同的位置计算候选模型的颜色直方图，并计算 Mean Shift 向量至收敛点；
- 4) 巴氏系数计算：在每次的 Mean Shift 迭代过程中，计算三种不同搜索窗的候选模型与目标模型之间的巴氏系数，经过不断地迭代，三种巴氏系数会达到稳定值；

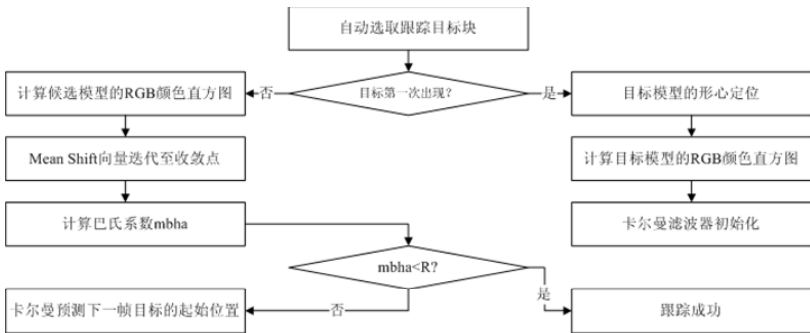


图 2 Mean Shift 和卡尔曼滤波相结合算法的实现流程设计

- 5) 取三种巴氏系数中的最大值 $mbha$ ，将其与一个较大的相似度门限 R 比较，若 $mbha < R$ ，说明目标模型与候选模型匹配，跟踪有效。
- 6) 把 Mean Shift 迭代算法获得的目标信息作为卡尔曼的测量向量，进行卡尔曼的更新，得到当前时刻状态向量的新估计，用于下一帧的迭代匹配。

4 算法优化

为了满足实时跟踪的需要，本文结合 C64x+ DSP 内核的特点，通过优化存储器的存取效率和提供程序的并行化程序来缩短程序运行所需要的指令周期数，加速代码的运行，对算法的实现进行了大量的编程方面的优化^[6]。具体的编程优化工作包含以下几个方面：

4.1 使用VLIB库

在优化过程中，本文采用了 TI 提供的 VLIB 库函数^[7]来对视频处理主线程中的混合高斯背景建模、腐蚀、膨胀、动目标检测、卡尔曼预测和更新等算法进行优化。VLIB 是 TI 公司针对视频处理 DSP 推出的高度优化的函数库，能方便地供用户使用，具有极高效的代码。表 1 是使用 VLIB 函数后对每个像素点处理所消耗的指令周期统计。由该表可以看出，采用 VLIB 函数，大大提高了程序的运行速度，为视频的实时处理提供了很好的基石。

表 1 VLIB 库的函数消耗的指令周期统计

算法中使用 VLIB 库的 API 函数	指令周期
VLIB_mixtureOfGaussiansS16	31.30 cycles/pixel
VLIB_dilate_bin_cross	0.27 cycles/pixel
VLIB_erode_bin_square	0.29 cycles/pixel
VLIB_createConnectedComponentsList	1.1-5.2 cycles/pixel
VLIB_kalmanFilter_2x4_Predict	154 cycles
VLIB_kalmanFilter_2x4_Correct	327 cycles
VLIB_convertUYVYint_to_YUVpl	0.4 cycles/pixel

4.2 IQmath函数库使用

TI 的 TMS320C64x+ IQmath 库是一些经过高度优化和高精度的数学运算函数的集合。C 程序员使用这个库可以把浮点运算无缝地移植到 TMS320C64x+设备上进行定点数运算。使用 IQmath 函数可以使你的代码执行速度比用标准的 ANSI C 语言编写的代码要快得多。同时，C 代码能够很方便地调用 IQmath 函数。在本实现中，很多的运算涉及到浮点运算，因此本文使用 IQmath 来编写高效的浮点数转定点数运算代码。

考虑到待处理的视频数据的范围和精度，本文优化时使用 Q15 的数据格式，这样既可以保证数据运算的不会溢出，也能达到比较高的精度。在代码优化过程，本文使用了_IQN()、_FtoIQN()、_IQNmpy()、_IQNsqr()、_IQNabs()等函数，大大地提高了运算能力，也解决了定点 DSP 上实现浮点运算的难题。

4.3 Cache的使用

由于高性能的 DSP 的片内高速存储资源有限，因此在系统实现时存储器的管理对提高整个系统的优化是非常重要的。为了满足算法的实时性，使用 Cache 来进行优化是很有必要的。

下面使用两种方法来解决算法实现时发生的 Cache 牺牲问题。

第一种方法是选择合适的数据类型，以减少内存需要。视频图像中很多数据都可以用 16 位来表示，这与 32 位相比可节省一半内存消耗，而且小数据类型容易实现汇编的 SIMD。

第二种方法是优化视频处理链。视频处理中的数据流是顺序的，即前一个函数的输出将作为下一个函数的输入。函数 1 的当前输入在 L1D 中，输出数据将被存放在低一级的存储空间（L2 或外部存储器），函数 2 在读取数据时会发生牺牲。为减少该类的 Cache 牺牲问题，可以将函数 1 的输出写入 L1D，则该部分数据可以直接重新访问而不会发生 CPU 停止。

4.4 内联函数的使用

C64x+编译器提供的内联函数可快速优化 C 代码。内联函数是直接映射为内联的 C64x+指令的特殊函数。在优化过程中，本文使用了许多的内联函数，比如_ADDSUB()、_MPY32U()等，提高了代码的执行效率。

5 实验结果与分析

本文在基于 DM6437EVM 系统平台上，并采用 PETS 2006 的标准测试视频源，利用 C 语言在 CCS 环境下编写了基于 DSP/BIOS 的 Mean Shift 与卡尔曼滤波结合算法实现程序，经过优化后，通过 Emulator 方式仿真实现了对地铁站内行人的实时跟踪。

测试视频源的图像分辨率为 720x576，帧率为每秒 25 帧，经过 TVP5146 后，视频处理帧大小为 640x480。算法的计算量与跟踪目标的个数、目标窗口的大小、Mean Shift 的最大迭代次数有关。目前算法实现只对单个行人进行全自动跟踪，选择 Mean Shift 的迭代次数为 10，由动目标检测算法提供目标模型的初始位置和目标块信息，采用 RGB（量化级为 8）空间的颜色直方图（特征值 bin 的个数为 512）建立目标模型和候选模型。图 3 给出了在 DM6437 上算法实现所得到的跟踪结果，图中白色矩形框为跟踪到的行人。如图所示，程序在第 5 帧自动检测出待跟踪目标并建立目标模型，在接下来的处理帧（25、40、60）中，系统能较好地跟踪到指定的行人。在实验过程中，比较拥挤的地铁站场景以及与目标颜色相近的物体会对跟踪结果造成较大的干扰，这主要是因为算法完全依赖于目标的颜色特征，未来的工作将考虑加入其他的特征信息来解决此类问题。经统计，优化后的程序能够达到 8 帧/秒的处理速度，DSP CPU 的负荷峰值为 55%，均值为 40%。



图 3 Mean Shift 与卡尔曼滤波结合算法在 DM6437 上的实现的跟踪结果

6 结论与展望

本文基于 DM6437 EVM 平台, 利用 TI 公司的 VLIB 函数和 IQmath 函数, 实现了结合 Mean Shift 和卡尔曼滤波的跟踪算法, 并且利用多种优化策略来提高算法效率, 满足了算法的实时、高效。实验表明在 DM6437 上实现的跟踪算法能对单目标进行良好地跟踪。未来工作将主要关注提高复杂场景下目标跟踪的鲁棒性以及扩展到同时对多目标进行跟踪。

参 考 文 献

- [1] Y.CHENG. Mean Shift, mode seeking and clusering[J].IEEE Trans. on Pattern Analysis and Machine Intelligence, 1995, 17(8): 790-799.
- [2] Chris Stauffer, W.E.L. Grimson, "Adaptive background mixture models for real time tracking," in CVPR, 1999.
- [3] Jie, Z., L. Wei-Jing, and S. Hui-Jia. A motion tracking method based on Kalman filter combined with mean-shift. in Wavelet Analysis and Pattern Recognition, 2008. ICWAPR '08. International Conference on. 2008.
- [4] TMS320DM6437 Digital Media Processor SPRS345L–NOVEMBER 2006–REVISED JUNE 2008
- [5] 彭启琮,管庆. DSP 集成开发环境—CCS 及 DSP/BIOS 的原理与应用. [M]北京:电子工业出版社, 2005 年
- [6] TMS320C6000 Optimizing Compiler v6.0 Beta, SPRU187N, July 2005.
- [7] Texas Instruments VLIB Application Programming Interface Reference Guide, SPRUG00A, April 2009.

作者简介:

赵渊 (1984-), 男, 湖南邵阳人, 硕士, 主要研究领域为信号处理;

管庆 (1968-), 硕士, 男, 副教授, 主要研究领域为信号处理、DSP 技术应用;

徐胜 (1978-), 男, 博士, 主要研究领域为图像识别。

一种新型格型IIR滤波器的研究与设计

傅文渊 凌朝东

(华侨大学 信息科学与工程学院 厦门 361000)

摘 要: 研究了格型 IIR 滤波器结构特点, 提出了一种新型的格型 IIR 滤波器结构。基于 DSP Builder 软件和 Altera 公司的 FPGA 芯片, 设计了所提出的新型格型 IIR 滤波器。通过 DSP Builder 和 Quartus II 的仿真分析, 新型的格型滤波器的占用逻辑资源减少了 51%, 器件工作的最高频率提高了 95%。所提出的格型滤波器能降低 FPGA 硬件资源利用率, 提高 IIR 滤波器的运算速度。

关键词: IIR; 格型滤波器; DSP Builder; FPGA

1 引言

在信号处理系统中, 数字滤波器是一项重要且普遍应用的技术。它通过一定的数学函数关系来改变输入滤波器信号中所含频率成分的相对比例或滤除某些频率成分。数字滤波器按其响应方式可以分为无限冲激响应滤波器(IIR Filter)与有限冲激响应滤波器(FIR Filter)两种^[1]。IIR 滤波器与 FIR 滤波器相比, 可用较低阶数获得较高的选择性, 在相同的时钟采样速率和相同的晶体管数量的前提下能提供更好的过渡带下降速率, 具有广泛的应用。

格型滤波器在 IIR 滤波器中起着重要作用。一是它的模块化结构便于实现数据高速并行处理, 易于 VLSI 集成; 二是一个 n 阶格型滤波器可以产生从 1 阶到 n 阶的 n 个横向滤波器的输出性能; 它对有限字长的舍入误差不灵敏。由于这些优点, 格型滤波器在现代谱估计、自适应滤波、语音处理、线性预测等方面应用广泛^[2]。

本文采用 FPGA 产商 Altera 公司开发的信号处理软件 DSP Builder 设计了一种新型格型 IIR 滤波器。此种格型滤波器较之传统的结构, 所占用的逻辑单元大幅度减少, 同时进一步提高格型 IIR 滤波器的运算速度。本文最后以 Altera 公司的 Stratix II 90 系列芯片作为 FPGA 的下载芯片, 利用 Quartus II 软件为开发平台, 进行了项目文件的分析、综合、编译适配、时序仿真分析和网表复写。根据最后的实验, 新型的 IIR 滤波器工作的信号频率由原来的 26.3MHz 提高到现在 52.3MHz, 所占用的逻辑单元数有原来的 372 个下降为 183 个。

2 传统格型滤波器的设计

本数字 IIR 滤波器的系统实现有各种结构形式。Gay 和 Marekl 于 1973 年提出了一种经典的格型滤波器结构。其基本单元的形状与快速傅立叶变换中的蝶形单元类似, 基本的数学理论是 Schur 算法。在格型滤波器中, 分母应用 Schur 算法来分析, 分子部分则用 Schur 算法对分母处理的结果为正规直交基函数的多项式展开算法进行分析。

稳定的 IIR 系统应满足以下条件:

$$\Omega_N(z)=\frac{z^{-1}\{\Phi_N\Omega_N(z)-\Phi_0\Omega_N(z)\}}{\sqrt{\Phi_N^2-\Phi_0^2}}=\frac{\Omega_N(z)-l_N\Omega_N^*(z)}{z\sqrt{1-l_N^2}} \tag{1}$$

$$\Phi_N=\sum_{i=1}^N\Phi_i z^i \tag{2}$$

$$l_N=\frac{\Phi_0}{\Phi_N} \tag{3}$$

由以上 3 式，可以得到格型滤波器的数学理论 Schur 算法^[3]具体实现。利用 $\Omega_N(z)$ 的表达式递推得到 $\Omega_{N-1}(z)$ 。由于两者的阶数相差为 1，因此，可以通过相同的方法循环得到从 $\Omega_N(z)$ 的所有 N 个表达式。并且这 N 个表达式中的常数项绝对值小于 z 最高次幂的的绝对值。

图 1 表示一个传统格型 IIR 滤波器的结构示意图。输入信号在每经过一个相同的模块，递归多项式的次数将降低一次。经过所有的 N 个模块后，递归多项式变为常熟。通过匹配这些模块的常数，就可以使用简单的乘法和延迟模块把这些串联起来，构成完整的 IIR 滤波器系统。

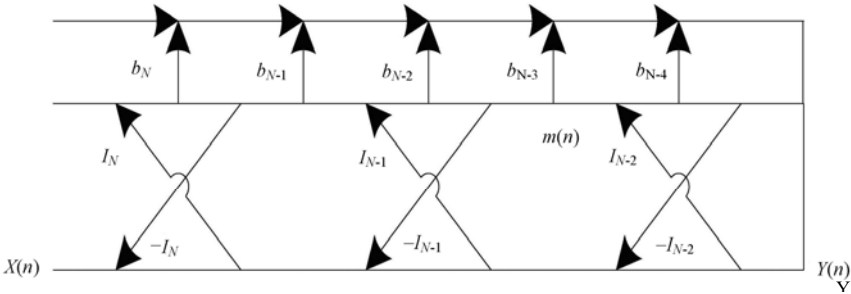


图 1 Schur 算法设计的格型 IIR 滤波器系统

3 新型格型IIR滤波器的设计

本文提出的新型格型 IIR 滤波器结构是基于传统 IIR 滤波器的结构。具体的结构图如图 2 所示。

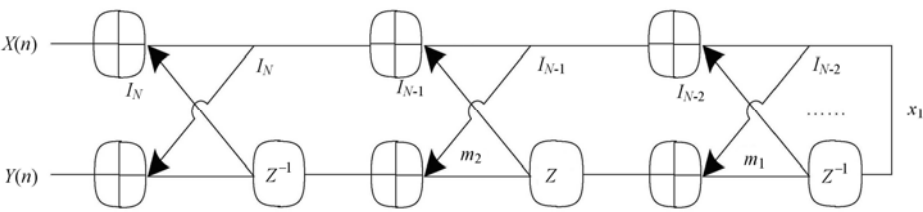


图 2 新型的格型 IIR 滤波器系统

图 2 示出了 3 阶格型 IIR 滤波器的结构图。由上图所示，为方便计算， N 阶格型 IIR 滤波器的最后一阶信号流入延迟单元的节点记为 x_1 ， x_1 前的信号节点记为 m_1 ，同理可得 m_2, m_3, \dots, m_N 。由信号系统的基本原理，可以得到以下等式，

$$x_1 z^{-1} = m_1 \tag{4}$$

由信号流向及递归思想，可以得到，

$$m_i = ((x_1 - \sum_{k=1}^{i-1} m_k l_k) l_i + m_{i-1}) z^{-1} \tag{5}$$

$$y(n) = (x_1 - \sum_{k=1}^{N-1} m_k l_k) l_N + m_{N-1} \tag{6}$$

$$x(n) = -m_{N-1} + (x_1 - \sum_{k=1}^{N-1} m_k l_k) l_N \tag{7}$$

因此，

$$H(z) = \frac{(x_1 - \sum_{k=1}^{N-1} m_k l_k) l_N + m_{N-1}}{-m_{N-1} + (x_1 - \sum_{k=1}^{N-1} m_k l_k) l_N} = \frac{1 + \sum_{k=1}^{N-1} z^{-k} a_k}{1 - \sum_{k=1}^{N-1} z^{-k} b_k} \tag{8}$$

其中 a_k, b_k 满足 Levinson-Durbin 递归的条件，具体的数值可以参考 Levinson-Durbin 算法来求得，读者可以参考文献^[4]，在此不再赘述。本文采用软件 DSP Builder 来仿真所提出的新型格型滤波器的性能，具体的电路实现为图 3。其中，仿真的信号源是 Matlab 软件中 Simulink 库的随机信号发生器、正弦信号发生器和两端口示波器。

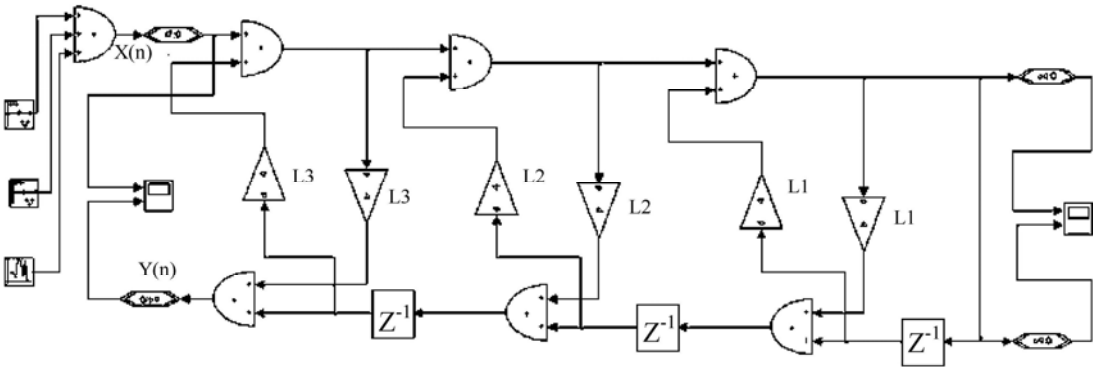


图 3 新型的格型 IIR 滤波器仿真电路图

4 滤波器电路的仿真

DSP Builder 在算法友好的开发环境中帮助设计人员生成 DSP 设计硬件表征，从而缩短了 DSP 设计周期。采用 DSP Builder 设计的第一步是在 Matlab 中的 Simulink 进行设计输入，即在 Simulink 的环境中建立 MDL 模型文件。用模块化的方式调用 Altera DSP Builder 和其他 Simulink 中的模块，构成系统级的仿真，具体的算法级仿真如图 3 所示。

输入格型 IIR 滤波器的信号采用频率分别为 98MHz、201MHz 的正弦信号与随机噪声信号的叠加，其中随机噪声信号的幅度为 4，所输入的正弦信号的幅度为 10。格型滤波器的采样频率为 489MHz，要求能够滤除高频信号和随机噪声信号。

仿真波形如图 3 所示，下方图是输入信号，上方图是输出信号。输入信号通过 IIR 格型滤波器后基本上变成单频的正弦信号输出。可见，该 IIR 滤波器可以比较好的消除随机噪声信号

和高频信号。

Matlab 中的 Simulink 对已经设计好的格型 IIR 滤波器系统进行编译,通过调用 DSP Builder 的 Signal Compiler 工具可直接生成 Quartus II 工程文件,再调用 Quartus II 完成综合,

网表生成和适配,直至完成 FPGA 的配置下载过程。本文选用的 FPGA 芯片为 Stratix II EP2S90F1020C4。仿真结果是,传统格型 IIR 滤波器所占用的逻辑单元数为 372 个,信号最高工作频率为 26.3MHz,滤波器的建立时间为 7.813ns,保持时间为 0.194ns。本文提出的新型格型 IIR 滤波器所占用的逻辑单元数为 183 个,信号最高工作频率为 58.3MHz,滤波器的建立时间为 7.708ns,保持时间为 0.233ns。

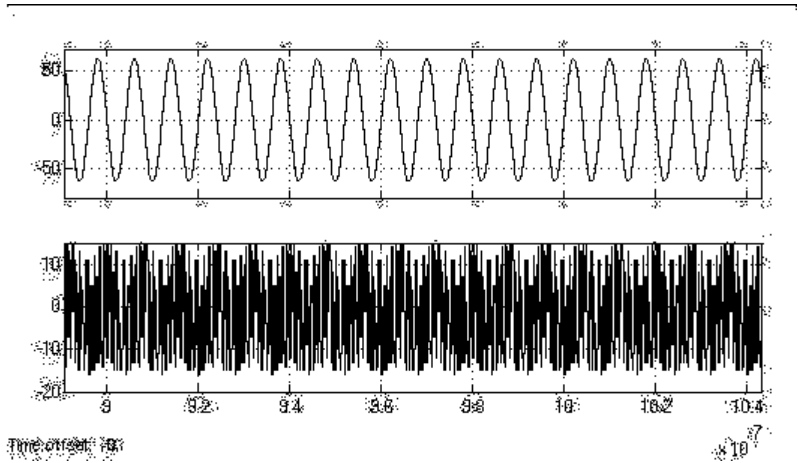


图 4 新型的格型 IIR 滤波器仿真波形图

由上面两组仿真数据对比可知,IIR 滤波器的建立时间、保持时间和关键路径延时相差不大。延时的大小与连线的长短和逻辑单元的数目有关,同时还受器件的制造工艺、工作电压、温度等条件的影响^[5]。因此,单纯的改变系统电路结构并不能根本的改善 IIR 滤波器的建立时间、保持时间等参数的大小。但新型格型 IIR 滤波器下载到 Stratix II EP2S90F1020C4 中所占用的逻辑单元数比传统的格型 IIR 滤波器减少 51%,器件工作的最高频率较之传统格型 IIR 滤波器提高了 95%。由此得出,新型格型 IIR 滤波器在占用资源数和工作速度等性能都有较大改善。

5 结论

采用格型 IIR 滤波器的数学理论,提出了一种新型的滤波器电路图。在此基础上,利用 DSP Builder 工具设计了 3 阶新型格型 IIR 滤波器。通过与传统格型 IIR 滤波器的仿真结果的对比得出,本文提出的格型 IIR 滤波器提高了信号的最高工作频率和减少所占用 FPGA 芯片的逻辑单元数等性能指标。本文的研究成果对利用 DSP Builder 进行数字信号处理和格型 IIR 滤波器的设计具有一定的参考价值。

参 考 文 献

- [1] 翟海涛,杨军,朱江. 一种基于 FPGA 的高速 FIR 滤波器的设计[J]. 信息化研究, 2009, 04.
- [2] R.W. Jones,B.L. Olsen,B.R. Mace. Comparison of convergence characteristics of adaptive IIR and FIR filters for active noise control in a duct. Applied Acoustics[J], 2007 68(7): 729 - 738
- [3] 彭煊, 杨红卫, 刘金福等, 基于 Schur 运算格的线性 ICA 估计算法[J], 电子学报, 2004, 03
- [4] Koukoutsis, Error sources and error propagation in the Levinson-Durbin algorithm[J], Signal Processing, IEEE Transactions on 1993 , 41(4): 1635 – 1651
- [5] Braslavsky, J.H., Middleton, R.H, Feedback Stabilization Over Signal-to-Noise Ratio Constrained Channels[J], Automatic Control, IEEE Transactions on 2007,52(8):1391-1403

作者简介:

傅文渊, 生于 1983 年, 男, 汉族, 福建省邵武人, 电子科技大学微电子与固体电子学院硕士毕业, 现工作于华侨大学信息科学与工程学院, 研究方向为高速数字信号处理与软件无线电前端电路设计等。

多视点图像的本征图分解

施柏鑫 李扬曦 许 超

(北京大学机器感知与智能教育部重点实验室, 北京, 100871)

摘 要: 本征图分解是计算机视觉中处理光照的一种重要手段, 然而对于多视点图像的本征图分解, 一直没有完整的解决方案。本文提出了一种多视点图像的本征图分解方案。通过输入不同光照下从不同视点拍摄的图像, 首先采用 SIFT 流的方法对多视点图像进行逐像素的匹配和对齐, 然后基于这些对齐的图像采用对同一视点的序列图像进行本征图分解的方法进行求解。实验结果表明, 这一套全新的处理流程能够准确地对多视点图像进行本征图分解。

关键词: 本征图分解; 多视点图像; SIFT 流

Intrinsic Image Decomposition for Multi-view Image

Shi Bo xin Li Yang xi Xu Chao

(Key Laboratory of Machine Perception (MOE), Peking University, Beijing, 100871, China)

Abstract: Intrinsic image decomposition is a critical pre-processing to deal with illumination in computer vision research. However, until recently, few works have been proposed to deal with the intrinsic image decomposition for multi-view images. In this paper, a novel workflow in solving this problem is introduced, given a set of multi-view images under different illuminations. SIFT flow is first adopted to perform a pixel-wise matching and alignment, then the aligned images can be decomposed following the similar procedure for image sequences of the same scene. The experiment results show that the proposed scheme can make this problem well solved.

Keywords: Intrinsic image decomposition; Multi-view image; SIFT flow

1 引言

本征图的概念是由 Barrow 首次提出的^[1]。在这篇经典的文章中, 本征图的分解定义为将一幅图像 I 分解为反射图 R (Reflectance) 和照射图 S (Shading), 记做 $I=RS$ 。其中反射图表示物体表面的本质颜色, 就是不受光照影响的部分; 而照射图可以反映物体表面法线方向及其光照的效果。这种分解可以作为很多和光照有关的计算机视觉问题的预处理操作, 为问题的进一步求解提供帮助。比如在一个视频监控系统中, 我们希望待分析的图像能够尽量避免由光照变化带来的误检, 此时如果使用本征图分解后的反射图作为监控系统的输入, 就可以避免光照影响; 另一方面, 由于照射图能够反映物体表面随光照的变化, 所以这种图像可以应用于根据光照恢复三维模型(例如 Shape from shading)的问题求解。

然而本征图的分解是计算机视觉中的一个难题。因为我们想要从一幅输入图像得到同样

尺寸的两幅输出图像，这显然是一个病态问题，需要额外的约束条件才能求解。Retinex 是最早提出的求解这一问题的算法之一^[2]。著名的 Retinex 假设主要是指：物体表面反射的变化对应于较快的图像梯度变化，而光照引起的变化对应的图像梯度变化较慢。尽管这一假设并不能适用于所有场合，但是它为后来很多改进算法的提出奠定了基础，比如文献[3]中的改进算法，就是基于 Retinex 在一定程度上解决了本征图分解的问题。Tappen 同样利用了这一假设^[4]，并且尝试利用一个经过训练的二值分类器来区别由反射和光照引起的梯度变化，实验室结果较好，但是分类器的训练过程比较复杂。所以到目前为止，单幅图像的本征图分解仍然是一个研究当中的问题。

为了增加约束条件，一些研究者尝试从另一个角度来解决这一问题，即使用视频序列中的多幅图像。比如在 Weiss 的论文中^[5]，通过捕捉一个固定视点下不同光照的图像（比如一天中不同时刻的同一个建筑物），他证明这种情况下的本征图分解是一个最大似然估计问题，并进行了成功的求解和验证。但是这种方法最大的问题在于需要场景固定的不同光照图像，在一定程度上限制了该方法的实用。

本文提出的多视点本征图分解是一个目前还很少涉及的研究课题。因为在日常生活中，我们对于同一目标进行图像的捕捉往往得到的是不同视点的图像，相对于文献[5]的输入条件限制，本文对于输入图像的要求进一步放宽，即可以使用不同视点、不同光照的多幅图像，但是这一问题的难度比文献[4]中的单幅图像问题略低，不过我们要求算法的复杂度也更低（不使用复杂的分类器）。我们首先通过最新提出的 SIFT 流算法^{[6][7]}对多视点图像进行逐像素的匹配和对齐，使其统一在同一个视点上；然后采用文献[5]中的算法进行本征图求解。实验结果证明使用这两种算法结合的全新处理流程，可以很好地解决多视点本征图分解的问题。

2 SIFT流算法简介

SIFT(Scale Invariant Feature Transform)特征^[8]是一种能够对图像的旋转和缩放保持不变，同时在一定程度上不随图像的光照和视角变化的特征，目前已经在很多的模式识别和分类算法中得到了广泛的应用。传统的图像匹配（比如视频编码中经典的块匹配）主要基于图像亮度的绝对值差，这种方法对于两幅存在亮度差异的图像存在很大的误差，从而不能得到正确的匹配。而且从不同角度拍摄的多视点图像经过了摄像机内部复杂的投影变换，它们之间的映射关系是不能通过简单的平移、旋转、甚至透视变换来描述的。因此，描述多视点图像之间的变换关系需要鲁棒的特征和复杂的映射函数。

直到最近，Liu 在他的两篇文章中^{[6][7]}巧妙地将 SIFT 特征用于稠密图像匹配问题，同时结合光流估计的优化过程，提出了 SIFT 流的概念。通过 SIFT 流，图像的每一点都是一个 128 维的 SIFT 特征，这一特征充分考虑了周围像素的情况而且具有多种的不变性，所以可以进行逐像素的匹配，即一一映射，实现把一个视点的图像变换到另一个视点，或者说用一个视点的图像重构出另一个视点。这一概念可以通俗地与光流进行类比：即光流是时间上的稠密匹配，而 SIFT 流是图像空间上的稠密匹配，可以实现多视点图像、甚至内容相似的不同场景图像的对齐。这种对齐，在传统的全局图像配准和块匹配下都是无法实现的。SIFT 流的求解可以用一个需要最优化求解的目标函数来描述：

$$E(\mathbf{w}) = \sum_{\mathbf{p}} \min(\|\mathbf{s}_1(\mathbf{p}) - \mathbf{s}_2(\mathbf{p} + \mathbf{w})\|_1, t) + \sum_{\mathbf{p}} \eta(|u(\mathbf{p})| + |v(\mathbf{p})|) + \sum_{(\mathbf{p}, \mathbf{q}) \in \mathcal{E}} \min(\alpha|u(\mathbf{p}) - u(\mathbf{q})|, d) + \min(\alpha|v(\mathbf{p}) - v(\mathbf{q})|, d)$$

在这一目标函数包括三项（一个求和对应一项）。数据项用于度量不同位置像素 SIFT 特征 $s(\mathbf{p})$ 的差异，其中的 $\mathbf{p} + \mathbf{w}$ 表示两幅图像像素的对应关系（类似于运动矢量）；随后的幅度项用于约束图像的匹配对应关系在一定的距离之内；最后的正则项用于平滑约束，使得相邻像素保持一定的关联。通过对这一目标函数的优化，能量最小时相应的逐像素对应关系就是 SIFT 流匹配的结果，即不同视点图像的稠密匹配。

上述最优化的求解是一个比较复杂过程，需要一些数据结构的优化才能用于图像这种大规模的矩阵运算，具体的解法请参考 Liu 在今年 CVPR 的最新文章[7]。

3 本征图分解算法

经过 SIFT 流匹配和对齐的多视点图像可以近似认为是同一个视点的图像，因为 SIFT 流可以不受光照变化影响，实现高精度、逐像素的稠密匹配。如果我们对于不同光照条件下的多视点图像使用 SIFT 流进行对齐，则这一预处理的结果使得文献[5]中算法的输入条件得到满足，从而可以进行本征图的分解。这一整套算法的流程如图 1 所示。

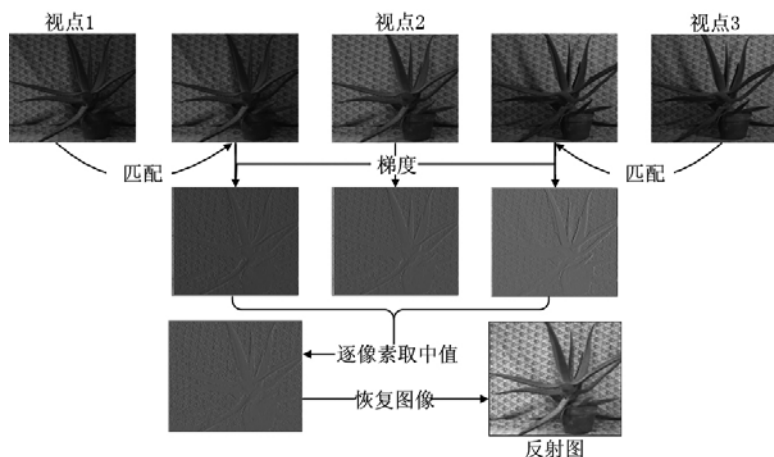


图 1 多视点图像本征图分解流程

算法的输入是不同视点、不同光照的多幅图像（这里以三幅为例）。我们指定视点 2 的图像为参考图像，首先通过 SIFT 流的匹配过程，将视点 1 和 3 的图像对齐到视点 2，这样得了三个视点逐像素对齐的图像。然后对这三幅图像分别取水平和垂直方向的梯度图，并对梯度图进行逐像素对应的取中值操作，得到一幅由序列中各梯度图的中值构成的图像（图 1 左下角）。这一图像中包含的所有像素的梯度均是只与反射有关的梯度变化，所以通过最后一步对这一图像的恢复（梯度的逆运算，即从梯度图像恢复原始图像）即可得到本征图分解后的反射图（图 1 右下角）。有了反射图以后，照射图可以直接求解（根据 $I=RS$ ）。这一过程的严格数学证明（主要说明取中值操作和反射图梯度图的等价关系）请参考文献[5]。从实现的角度，

这一流程中的取梯度和中值都是基本的图像处理操作；最后一步的恢复过程是类似积分的运算，可以通过求解泊松方程得到结果，也可以采用文献[5]中的解卷积方法。具体过程不再赘述。

4 实验结果

我们采用了 Middlebury 的立体视觉数据集^[9]中的部分图像对这一全新的处理流程进行验证。这一数据集中包含了多组多视点的图像，同时每个视点又在不同的光照条件下进行采集。我们从中选取了两组图片，每组 3 个视点，每个视点一种不同的光照条件。根据图 1 的流程，我们首先将视点 1 和 3 对齐到视点 2，然后进行本征图的分解。图 2 的实验结果中，(1)和(3)行分别代表不同的实验图像组，(2)和(4)行的(a)-(c)列为分解得到的对应视点的照射图，而(2d)和(4d)为反射图。可以看到，每个照射图反映了一种不同的光照配置，光源方向的变化可以明显地解读；而反射图保留了物体表面不随光照变化的部分，比如(4d)左上方纸上的涂鸦线条就被正确地保留在了反射图中，而没有显示在照射图中，因为这些笔划的痕迹是不会随光照改变的。图 2 中的(1d)和(3d)是一个简单的对比实验，显示了通过全局配准（平移+选转）对齐再进行本征图分解得到的反射图结果。可以看到这两幅图中近处物体的边缘存在很明显的虚化现象（照射图中亦存在同样的问题，限于篇幅，结果从略），这一结果是很不准确的。这更证明了多视点图像的匹配和对齐不是一个简单的变换和映射，很难使用已有的视点变换算法进行对齐，而本文采用的 SIFT 流算法能够进行准确的稠密匹配，从而可以得到清晰的本征图分解结果。

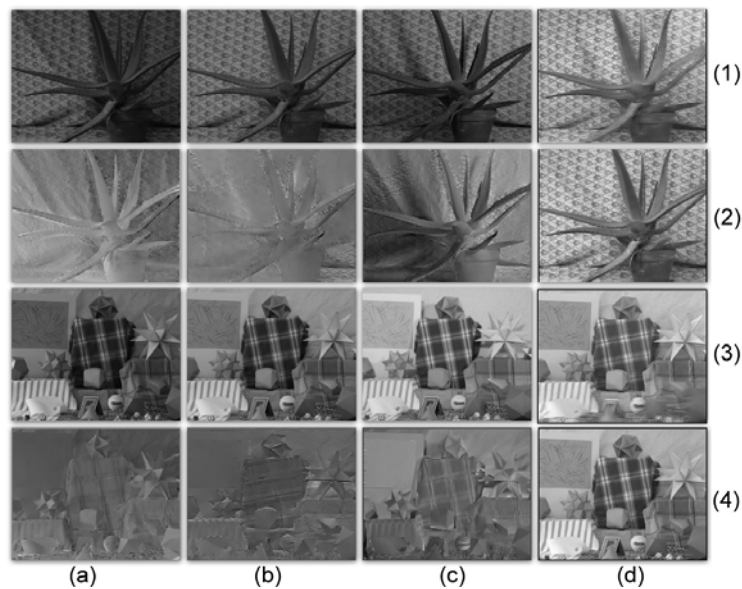


图 2 多视点本征图分解结果

5 总结

本文通过 SIFT 流对多视点图像对齐得到稠密匹配图像，然后利用多幅同一视点、不同光

照图像的本征图分解算法，成功地对不同光照的多视点图像进行了本征图分解，使得本征图分解可以使用更容易得到的多视点图像作为输入。实验结果证明，分解的图像能够很好的区分物体表面的反射和光照引起的变化，可以用于帮助解决其他的计算机视觉问题。然而目前的结果只是验证了方法的可行性，随着下一步工作的深入，更多样的数据测试和更完善的结果分析将陆续完成。

参 考 文 献

- [1] BARROW H. and TENENBAUM J., “Recovering intrinsic scene characteristics from images. In Computer Vision Systems” [M], A. Hanson and E. Riseman editors Academic Press, 1978.
- [2] LAND E. and MCCANN J., “Lightness and retinex theory” [J], Journal of Optical Society of America, Vol. (3): 1684–1692, 1971.
- [3] KIMMEL R. and et al., “A variational framework for retinex” [J], IJCV, Vol. (52): 7–23, 2003.
- [4] TAPPEN M., FREEMAN W. and ADELSON E., “Recovering reflectance from a single image” [J], IEEE Trans. PAMI, Vol. 27(9): 1459–1472, 2005.
- [5] WEISS Y., “Deriving intrinsic images from image sequences” [C], In Proc. ICCV, 2001.
- [6] LIU C. and et al., “SIFT flow: dense Correspondence across different scenes” [C], In Proc. ECCV, 2008.
- [7] LIU C., YUEN J. and TORRABLA A., “Nonparametric scene parsing: label transfer via dense scene alignment” [C], In Proc. IEEE CVPR, 2009.
- [8] LOWE D., “Object recognition from local scale-invariant features” [C], In Proc. IEEE ICCV, 1999.
- [9] HIRSCHMULLER H. and SCHARSTEIN D., “Evaluation of cost functions for stereo matching” [C], In Proc. IEEE CVPR, 2007.

基于行移字移编码算法的文本水印鲁棒性研究

黄兴莉 邵怀宗 彭启琮

(电子科技大学 通信学院, 四川 成都 610054)

摘要: 二值文本的特殊性决定了其水印技术的特殊性。在该文中,根据文本的特点,在原有的行移字移编码算法基础上获得了一种改进算法即间距检测法,并进行了 MATLAB 的仿真和分析,结果表明改进的算法鲁棒性能好。

关键词: 行移字移编码; 行间距; 字间距; 鲁棒性; 边界信息

Research on a Robust Text Watermarking Algorithm Based on Line Shifting and Word Shifting

Huang Xing li Shao Huai zong Peng Qi cong

(College of communication, University Of Electronic Science And Technology Of China, Chengdu
Sichuan 610054, China)

Abstract: The peculiarity of one text image has decided the particularity of the watermarking technology. In this paper, based on line shifting and word shifting, an amending algorithm named space detection method has been proposed. At the same time, some experiments were done to analyze the performance. The results show that the robustness of this amending algorithm is better than former one.

Keywords: line shifting and word shifting coding; line space; word space; robustness; boundary information

1 引言

文本图像是一类特殊的静止图像,它是二值的,整个图像只包含“黑”和“白”两种颜色。像素值是单一的“0”和“1”。对文本中任何一个像素值进行改变都会造成严重的视觉失真。因此传统的适用于普通图像的如 LSB, DCT, DWT 等数字水印算法对于文本图像都不适合^[1]。但是文本也有自己的特点,比较显著一点就是文本都是由一行行的字有规律的分布组成的,行和行之间以及字和字之间都有一定的空隙;同时,行之间的空隙以及行的高度大致相等,字和字之间的空隙也大致相等。基于文本中的这个特点,可以在其垂直方向上采取行移编码算法^{[2]-[3]},在水平方向上采取字移编码算法^{[2]-[3]}。在本文中,结合文本的特点在行移字移编码算法的基础上提出一种新的检测方法——间距检测法,并结合水印比特流的编码操作提高了整个系统的性能。大量的 MATLAB 的实验表明,改进的行移字移编码算法原理简单,计算量很小,鲁棒性能好。

2 改进的行移字移编码算法

2.1 水印的嵌入

当某一行及其两个相邻行都足够长，而且相邻行之间的距离均相等的情况下该行方可嵌入水印。可以选择偶数行作为控制行保持不动,而奇数行则按照水印比特流的要求进行上下移动嵌入水印。控制行的作用是失真补偿和估计剩余噪声^[4]。

字移编码算法和行移编码算法相似，只是对字的定位是在对行的定位基础上进行的，水印比特流改变的是字的左右边界。现在假设垂直方向上和水平方向上需要嵌入的水印分别是: $[0\ 1\ 0\ 1\ 1\ 0\ 0]$ 和如图 1 所示的“黄”字（ 13×13 ），则原始图像图 2 经过水印嵌入系统之后的仿真结果如图 3 所示：



图 1 水平方向上需要嵌入的水

对数字水印技术研究兴趣的突然增长源于人们对版权问题的关注。近年来，随着计算机多媒体的迅猛发展，人们可以方便地利用数字设备制作、处理和存储图像、语言、文本和视频等信息媒体。与此同时，数字网络通信正在飞速发展，使得信息的发布和传输实现了“数字化”和“网络化”。数字媒体很容易借助互联网播或者随机储存被复制、处理、传播和公开。这样就引发出数字信息传输的安全问题和数字产品的版权问题。为了解决信息安全和版权保护问题，数字产品所有者首先想到的是加密和数字签名等技术。但这些方法在实际应用中出现诸多不足，因此需要寻求一种不同于传统技术的更加有效的手段，来保障数字信息的安全传输和保护数字产品的版权。数字水印技术能够使嵌入在数字产品中的信息不会被常规处理操作去除。数字水印技术一方面弥补了密码技术的缺陷，因为它可以为解密之后的数据信息提供更进一步的保护，另一方面，它也弥补了数字签名技术的缺陷，因为它可以在原始数据中一次性嵌入大量的秘密信息。

图 2 原始的文本图像

对数字水印技术研究兴趣的突然增长源于人们对版权问题的关注。近年来，随着计算机多媒体的迅猛发展，人们可以方便地利用数字设备制作、处理和存储图像、语言、文本和视频等信息媒体。与此同时，数字网络通信正在飞速发展，使得信息的发布和传输实现了“数字化”和“网络化”。数字媒体很容易借助互联网播或者随机储存被复制、处理、传播和公开。这样就引发出数字信息传输的安全问题和数字产品的版权问题。为了解决信息安全和版权保护问题，数字产品所有者首先想到的是加密和数字签名等技术。但这些方法在实际应用中出现诸多不足，因此需要寻求一种不同于传统技术的更加有效的手段，来保障数字信息的安全传输和保护数字产品的版权。数字水印技术能够使嵌入在数字产品中的信息不会被常规处理操作去除。数字水印技术一方面弥补了密码技术的缺陷，因为它可以为解密之后的数据信息提供更进一步的保护，另一方面，它也弥补了数字签名技术的缺陷，因为它可以在原始数据中一次性嵌入大量的秘密信息。

图 3 嵌入水印之后的文本图像

2.2 水印的提取

水印的提取与嵌入是一个相反的过程，但要补偿文本由于衰减，平移，旋转，缩放，噪声攻击等操作所引起的失真。本文中我们采用一种新的提取水印的方法——间距检测法。

采用改变行间距字间距的方法来嵌入水印的文本有一个显著的特点，那就是无论嵌入的水印是“0”还是“1”，只要首行和每一行的首字保持不动，那么嵌入水印之后总的行间距之和以及每一行上总的字间距之和是相等的，唯一不同的只是行间距和字间距与平均行间距以

及平均字间距之间的大小关系不一样。文本行与行之间的间距 space 可以用如下的函数表示 (其中定义 hang 为文本的行数, row 为文本的像素行数, “1” 表示上移, “0” 表示下移):

$$\text{space}(i) = f(i, j); i \in (1, \text{hang} - 1), j \in (1, \text{row}) \tag{1}$$

那么, 无论嵌入的水印比特流是怎样分布的都有 (2):

$$\sum_{i=1}^{i=\text{hang}-1} \text{space}(i) = \text{固定的常数} \tag{2}$$

则每一行的相对行间距 det 如 (3)所示:

$$\text{det}(i) = \text{space}(i) - \sum_{i=1}^{i=\text{hang}-1} \text{sapce}(i) / (\text{hang} - 1), i \in (1, \text{hang} - 1) \tag{3}$$

相对行间距必然会出现交叉的正负关系。水印信息就是根据这样的正负关系来提取, 水印信息 water 的提取如(4)所示:

$$\text{water}((i + 1)/2) = \begin{cases} 0 & \text{det}(i) > 0 \ \& \ \text{det}(i) < 0 \\ \square\square\square\square\square & \text{det}(i) = 0 \\ 1 & \text{det}(i) < 0 \ \& \ \text{det}(i) > 0 \end{cases} \tag{4}$$

字移解码和行移解码是相同的, 只是字移解码比较的是字间距之间的关系, 它和行间距具有同样的特点。采用此方法在本实验中提取到的水印分别如图 4 所示:

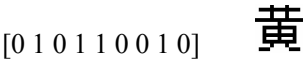


图 4 间距检测法提取到的水印

3 行移字移编码算法的鲁棒性研究

任何一种数字水印算法都要求其对于有意或者无意的攻击具有一定的鲁棒性, 文本水印也不例外^[5]。对于本文的行移字移编码算法而言, 其嵌入和提取水印的关键都在于对边界信息的确定, 边界信息包括行的上下边界以及字的上下左右边界, 而与文本中每一个字的具体形状无关。从式(4)我们可以看出, 水印的提取只是比较间距相对于平均间距的大小关系, 因此这种比较可以允许间距在一个比较大的范围内变化, 只要满足正确的大小关系而不用满足固定的数学关系就可以提取出水印。从这个意义上说此算法能够提高系统的鲁棒性。在水印嵌入之前, 对原始的水印比特流采取了编码操作 (本文选取了差分编码方式), 把水印比特流以一种更加隐蔽和安全的方式嵌入到文本之中, 这样整个系统的鲁棒性能以及安全性都得到了较大的改善。为了说明其鲁棒性能, 本文用几个 MATLAB 实验加以说明 (以下的实验行移解码都是 100%正确, 关键是字移解码的差别, 因此行移解码结果就不再展示)。

3.1 旋转实验

旋转之后的图像读入时行的边界信息会发生变化, 行的错误定位会影响到对字的定位。行移字移编码可以检测出图像是否出现小幅度的旋转, 并能检测出图像是顺时针旋转还是逆时针旋转, 从而可以修正提取出来的水印信息。图 5 所示的水印依次是检测系统从顺时针旋转 0.5, 1, 1.5 度以及逆时针旋转 0.5, 1, 1.5 度之后的文本图像中提取到的。

3.2 噪声攻击实验

遭受噪声攻击之后的图像黑白像素分布的比例将会发生改变，某些噪声可能只是改变灰度值，但不会影响二值化之后的结果，但是某些噪声会改变字的边界信息。对于改变灰度值的噪声，行移字移编码算法具有很好的鲁棒性，对改变边界信息的噪声，此算法可以检测出噪声类型，排除部分干扰。图 6 所示的水印依次是检测系统从遭受密度为 0.5，0.4，0.2 的高斯噪声以及密度为 0.015，0.01，0.005 的椒盐噪声攻击后的文本图像中提取到的。

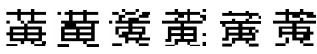


图 5 旋转后的文本图像通过检测系统提取到的水印

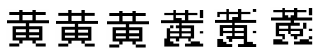


图 6 噪声攻击后的文本图像通过检测系统提取到的水印

3.3 文字篡改

传输的文本内容有可能被恶意篡改，不同的篡改方式对文字的边界信息影响不同。总的来说，由于中文汉字的行高是大致相等的，因此篡改对行移正确率影响很小，而不同的字的边界有可能不同，因此文字的篡改将会对字移解码产生一定的影响。图 7 所示的水印依次是检测系统从第三行和第六行的文字交换后的文本，第七行文字被其他文字替换后的文本，“可以”和“能够”交换后的文本，“安全”用“safe”替换后的文本中提取到的。

3.4 平移实验

图像在扫描的时候经常会发生平移现象，其文本的像素分布也会发生相应的平移。对于行移字移编码算法，由于平移是整个像素的平移，并没有改变行和字的边界信息，因此不会影响提取到的水印信息。图 8 所示的水印依次是检测系统从向左，向右，向上，向下平移 10 个像素之后的文本图像中提取到的。

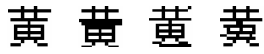


图 7 篡改文字的文本通过检测系统后提取到的水印

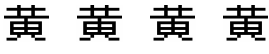


图 8 平移之后的图像通过检测系统后提取到的水印

3.5 剪切实验

剪切的位置以及剪切面的大小将会严重影响边界信息的确定。非敏感区域的剪切对水印提取的影响较小。敏感区域的剪切对解码有一定的影响，但是只要剪切部分不是足够大，仍然可以辨别出嵌入的水印信息。图 9 是检测系统从遭受不同剪切后的文本图像中提取到的水印。

亮度变化, 格式转换, 遭受滤波器的攻击等对行移字移解码都不会产生任何影响, 这里就不举例说明了。实际上, 加大攻击的强度和广度都不会影响到行移解码的检测结果。我们可以把最重要的信息以行移编码的方式嵌入到文本中, 而相关性强的以字移编码的方式嵌入, 同时也可以利用行移解码的结果来修正字移解码的结果, 实现系统较高的鲁棒性能。

4 总结

本文结合间距检测法对原有的行移字移编码算法进行了改进。间距检测法原理简单, 计算量很小, 并且相对间距在一个比较广的误差范围内都可以提取出正确的水印信息, 系统的鲁棒性能得以增加, 这是其他文本水印技术所不具备的优点。在水印嵌入之前采取了置乱以及编码技术, 把水印比特流以一种更隐蔽更安全的方式嵌入到文本中, 二者的结合不仅使得系统鲁棒性能提高而且使得嵌入的水印更加安全。本文还对此算法的鲁棒性进行了实验研究, 实验表明改进后的算法具有较强的鲁棒性, 尤其是行移编码算法, 字移编码算法鲁棒性能较弱但是水印嵌入量较大, 行移编码与字移编码的结合可以改善文本水印鲁棒性能不佳, 可嵌入水印量小的缺点。但是由于文字自身的特点, 行移字移编码算法还是不能使二者达到最佳, 因此需要继续寻找更佳的水印算法。

参 考 文 献

[1] 黄华, 齐春, 李俊, 等. 文本数字水印[J]. 中文信息学报, 2001, 5 月, p52-56.

[2] S.Low,N.Maxemchuk,J.brassil and L.O’Gorman.Document Marking and Identification using Both Line and Word Shifting[S].Proceedings of the IEEE.INFOCOM’95,Boston MA,April,1995,2:853-860

[3] J.Brassil,S.Low,N.Maxemchuk and L.O’Gorman.Electionic Marking and Identification Techniques to Discourage Document Copying(J).Proceedings of the Infocom’94,june,1994, p1278-1287.

[4] 杨洁.数字文本图像水印嵌入和提取算法研究[D].西安科技大学, 2006 年.

[5] 张黎.基于小波变换的二值图像技术[D].山东师范大学, 2007 年.

作者简介:

黄兴莉(1986-),女,四川内江市,硕士研究生,电子科技大学通信与信息工程学院,主要研究方向:现代通信中的信号处理;

邵怀宗(1969-),男,四川巴中市,博士,电子科技大学通信与信息工程学院,副教授,研究生导师,主要研究方向:无线通信及其信号处理,阵列信号处理和 DSP 技术与应用等;

彭启琮(1946-),男,四川乐山市,博士,电子科技大学通信与信息工程学院,教授,博士生导师,主要研究方向:通信与信息系统中的信号处理,高速信号处理与实时信号处理, DSP 技术等。

随机密钥预分布的抗攻击模型研究

张 啸 王 晖 林晓辉

(深圳大学 信息工程学院, 深圳 广东 518060)

摘 要: 传感器网络中由于非对称加密体制不适用于传感器节点, 而且节点易受物理攻击, 其密钥分配在一直是一个具有挑战性的问题。在基于随机密钥预分配的方案的无线传感器网络中, 网络的安全性和连通性始终是一对矛盾的统一体, 因此设计密钥池和密钥环的大小是至关重要的问题。这篇文章提出一种调节²密钥环大小的方法, 该方法能在保证全网安全连通性的前提下, 有效的提高网络的安全度。

关键词: 无线传感器网络; 安全; 随机密钥; 密钥管理

Random Key Establishment in Sensor Networks Using Anti-attack Model

Zhang Xiao Wang Hui Lin Xiao hui

(Institute of Information Engineering, Shenzhen University, Shenzhen, Guangdong 518060, China)

Abstract: Key establishment in sensor networks is a challenging problem because asymmetric key cryptosystems are unsuitable for resource constrained sensor network, in which, nodes are also vulnerable to physical compromise by an adversary. This problem is aggregated by the fact that, in a random key pre-distribution network, security is always in conflict with network connectivity. Thus, the design of key pool and key ring is critical for the network security. In this paper, by adjusting the key ring size, we present a method which can effectively improve the network resilience to hostile attacks, while the network connectivity can still be satisfied.

Keywords: Wireless sensor networks; Security; Random key; Key management

1 介绍

无线传感器网络已经被广泛的用于实时交通监测、军事感知和跟踪、野外探测等很多领域^[1]。分布式的传感器网络和 ad-hoc 网络包含成千上万个计算能力、存储能力和通信能力有限的传感器节点, 这些节点的分布的环境有可能比较恶劣, 通信过程时断时续^[2]。无线传感器网络的拓扑经常是动态的, 需要在布置及初始化后增加或删除节点。除此之外, 节点也有可能布置在敌方区域, 极易受到敌方的攻击, 再加上有限的计算能力、存储能力和通信能力^[3], 如何实施安全管理, 把一个传感器网络从一堆分立的节点, 或者说一个完全裸露的网络, 通

过预分配的形式使得网络能够安全的共享保密信息是个比较困难的问题^[4]。

为了解决无线传感器网络安全引导的核心问题,安全密钥的建立过程^[5], Eschenauer^[6]等人最早提出了随机密钥预分配模型,该模型是一个概率模型,其网络安全保障采用简单的共享密钥发现机制和路径密钥生成机制来实施。其思想是在一个比较大的密钥池中,每个节点都有密钥池中的一小部分密钥,只要节点之间拥有一对相同的密钥就可以建立安全的通道。其模型分三个阶段,密钥预分配阶段中由基站或高性能汇聚节点产生大密钥池及密钥的 ID,然后每个节点收到由基站从密钥池中随机抽取 m 个密钥,同时包括密钥的标识号,这 m 个密钥组成一个密钥环。共享密钥阶段中,一旦节点收到预分配的密钥后,节点将会布置在需要的场所,每个节点将寻找邻居节点中的共享密钥。密钥发现的方法有很多,最简单的是某个节点向它邻居节点广播它的密钥 ID,如果发现邻居节点有相同的密钥 ID,则以此密钥作为它们相互的通信密钥。当然广播交换 ID 的弊病是敌人可以通过交换 ID 分析出安全的网络拓扑,从而对网络造成威胁。另一种方法是:对于每个节点的 k 个密钥,发送节点可以向它的邻居广播一个列表, $\{\alpha, E_{k_i}(\alpha), i=1,2,3,\dots,m\}$, 其中 α 称为一个挑战。如果接收节点能够解密 $E_{k_i}(\alpha)$ 的话,那么就能将解密的那个密钥作为他们的共享密钥。路径密钥建立阶段中,当两个节点之间存在相同的共享密钥,他们可以建立一条安全的链路。如果他们之间没有共享密钥,那么这两个节点通过寻找第三方节点来寻求路径建立,如果第三方拥有和这两个节点相同的共享密钥,则由其作为该两个节点的临时的密钥分发中心,来分配其相同的密钥。

随后 Chan^[7]等人在随机密钥预分配的基础上提出了 q -composite 模型,基本模型中任何两个邻居节点的密钥环中至少需要一个公共密钥建立安全链接,而该模型将公共密钥的个数提高到 q ,从而大大加强了小范围攻击的安全性。但是相对于基本模型,要想安全网络中任意两点之间的安全连通度达到理想的概率,就必须缩小整个密钥池的大小,增加节点间共享密钥的交叠度。

所有无线传感器网络基于随机密钥预分配的方案中,安全和连通性始终是一对矛盾,这是因为,一方面在给定大小的密钥池下,要达到较高的连通性,就必须增加节点中存放密钥的数量,使节点之间共享密钥的概率增加,但是这样做的后果是一旦节点被敌方捕获,被获取的密钥数也会增多,因而会引起网络安全性的降低。另一方面,为了达到一定的安全性,希望敌方在俘获一定数量的节点而不能获取过多的密钥时,就必须尽可能地减少存放在节点中的密钥数量,结果就会降低节点之间共享密钥的概率,从而导致网络安全连通性的降低。因此如何设计密钥池和密钥环的大小成为了本文研究的关键。我们下文提出一种调节密钥环大小的方法,在保证全网安全连通性的前提下,提高网络的安全度,接下来我们用此方法依次分析基本随机密钥预分布模型和 q -composite 模型。

2 研究模型和方法

首先,网络通信连通度的分析基于一个随机图^[8,9] $G(n, pl)$, 其中 n 为节点个数, pl 是相邻节点之间能够建立安全链路的概率。根据 Erdos 和 Renyi 对于具有单调性的图 $G(n, pl)$ 的分析,有可能为图中的顶点计算出一个理想的连通度 d (它表示图中连接某个顶点的边的平均条数),使得图的连通率非常高,达到一个指定的门限 c (例如 $c = 0.999$)。Eschenauer 和 Gligor^[3] 给出的规模为 n 的网络的节点的理想连通度 d 如下式:

$$d = \frac{n-1}{n} (\ln(n) - \ln(-\ln(c))) \quad (1)$$

对于一个给定密度的传感器网络，假设 n' 是节点通信半径内邻居个数的期望值，则成功完成密钥建立阶段的概率可以表示为：(其中 n 为网络中的节点个数， n' 为邻居节点个数的平均值)

$$p = \frac{d}{n'} \quad (2)$$

2.1 基本随机密钥分布模型的抗攻击研究

我们的随机密钥预分布模型的初始化同样也是有三个阶段^[6,7]，密钥预分配阶段、共享密钥发现阶段和路径密钥实施阶段。在密钥预分配阶段，首先先生成一个大密钥池 S ，然后给每个节点随机选取 m 个密钥组成节点的密钥环，同时节点保存其密钥的 ID。在共享密钥发现阶段中，每个节点寻找邻居节点是否有相同的共享密钥，如果存在相同的密钥，则可以建立安全的链接，如果没有相同的密钥，则可以在路径密钥发现阶段，通过寻找第三方节点来寻求路径建立，如果第三方拥有和这两个节点相同的共享密钥，则由其作为该两个节点的临时的密钥分发中心，来给他们分配相同的密钥。

当任意两个节点的密钥环存在共同一个密钥时，他们之间能够建立的一条安全链路。考虑到密钥池的大小为 $|S|$ ，一个节点中密钥环的大小是 m 。我们可以计算出任意两个节点在密钥环中共享一个共同的密钥的概率 p 如下[6]：

$$p = 1 - \frac{(1 - \frac{m}{|S|})^{2(|S| - m + \frac{1}{2})}}{(1 - \frac{2m}{|S|})^{(|S| - 2m + \frac{1}{2})}} \quad (3)$$

接下来，我们评估网络的适应度 R ，依据的理由是当部分节点被捕获后，攻击者能够从被捕获节点中完全恢复出其密钥环，依此来计算网络通信信息可能被捕获的概率[10]。我们定义网络通信信息被捕获率的方程 f ，也就是当 x 个节点被捕获后，正常网络节点通信信息可能被捕获的概率。

设被捕获节点数为 x ，网络适应度 R 为当 x 个节点被捕获后，正常网络通信不受影响的概率

$$R = (1 - \frac{m}{|S|})^x \quad (4)$$

那么网络通信信息可能被捕获的概率的方程为

$$f = 1 - R = 1 - (1 - \frac{m}{|S|})^x \quad (5)$$

从式(3)和式(4)我们可以看出，通过调节 m 的大小，我们可以权衡 p 和 R 。本文提出的方案中，我们想给定一个理想的网络适应度 R_0 值，通过调节 m 的大小，以便得到较强的网络适应度 R ，然而我们可能要在一定范围内降低节点间的连通率 p ，根据公式(1)和(2)可以得出全网的连通性要降低。如图 1 所示，如果密钥环 m 减小，网络的适应度 R 将增强。在这个仿真中。密钥池的大小 $|S| = 100000$ 。

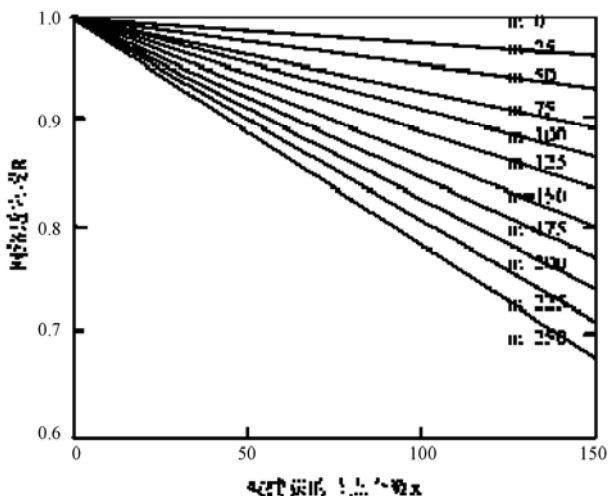


图1 不同的密钥环大小 m ，随着节点被捕获的个数 x 的增加，网络适应度的变化

前面我们已经提到，对于给定的节点的连通率 p ，我们希望减少密钥环的大小 m ，从而增加网络的适应度 R 。然而，网络的适应度 R 和两个节点共享一个密钥的概率 p 之间存在权衡。随后，我们将讨论节点的连通率 p 和一个节点中的密钥环的大小 m 之间的关系，通过仿真结果，我们可以证明这个权衡是可以接受的。

我们首先提出一个理想的适应度 R_0 ，该理想适应度和我们所处的环境和所承载的应用的安全性有直接关系。如果该场所要求安全性强，那么我们在设 R_0 的时候要尽可能的高。 R 和 R_0 值的范围都在 0 和 1 之间，我们想要找到一个 m ，使得在捕获 x 个节点以后，网络的连通率 R 大于 R_0 ，即

$$R = \left(1 - \frac{m}{|S|}\right)^x \geq R_0 \quad (6)$$

则，

$$m \leq |S| \left(1 - R_0^{\frac{1}{x}}\right) \quad (7)$$

因此，我们可以得到 m 的最大值，

$$m_{\max} = |S| \left(1 - R_0^{\frac{1}{x}}\right) \quad (8)$$

为了保证网络的适应度控制在理想范围， m 的值不能超过 m_{\max} 。设 m_0 为密钥初始化阶段密钥环的初始大小； p 为节点间的连通率，和 m 的取值正相关；同时密钥池的大小 $|S|$ 是预先给定的。例如，如果设定式 (3) 中 $p = 0.33$ ， $|S| = 100000$ ， $m_0 = 200$ ， $R_0 = 0.8$ 。如果 m 小于 m_{\max} 将确保网络的适应度大于或等于 R_0 。但是，如果密钥环的大小 m 较小，会导致网络连通率 p 降低。为了更好的解决适应度和节点间连通率之间权衡的问题，我们提出了下面的方案，先设定密钥环的大小 m' 为变量，由于 m' 的变化，引起适应度 R 和两节点间连通率 p 发生变化。

$$R' = \left(1 - \frac{m'}{|S|}\right)^x \quad (9)$$

$$p' = 1 - \frac{(1 - \frac{m'}{|S|})^{2(|S| - m' + \frac{1}{2})}}{(1 - \frac{2m'}{|S|})^{(|S| - 2m' + \frac{1}{2})}} \quad (10)$$

如果 ($m_{\max} < m_0$) , 则

$$m' = m_{\max} \quad (11)$$

式 (10) 中, 如果 ($p' \leq p_{\min}$) , 使

$$m' = m_{\min} \quad (12)$$

其中 p_{\min} 为节点间连通率的最小阈值, 该阈值的作用是为保证全网的连通率可以正常通信。

如果 ($m_{\max} \geq m_0$) , 则使得

$$m' = m_0 \quad (13)$$

在本文的仿真中, 我们给定 $p_{\min} = 0.6p$, 例如, 如果 $m_0 = 200$, $|S| = 100000$, $p = 0.33$, $R_0 = 0.8$, 则 $p_{\min} = 0.2$, $m_{\min} = 160$ 。同样我们可以定义函数 f'

$$f' = 1 - R' = 1 - (1 - \frac{m'}{|S|})^x \quad (14)$$

需要说明一点, 从式(11)、(12)、(13) 中不难发现该模型中节点的密钥环在遇到攻击后会减小, 因此涉及到如何删除密钥环中密钥的问题, 我们采用的方法是随机删除密钥环中的密钥, 这样部分链路可能需要重新进行路径密钥的发现过程。

2.2 q-composite随机密钥预分布模型的抗攻击研究

q-composite 随机密钥预分布模型[7]和基本模型的过程相似, 只是要求相邻节点的公共密钥数要大于等于 q , 而且在建立共享密钥时用实际的公共密钥数 q' ($q' > q$) , 来建立共享密钥 k , $k = \text{Hash}(k_1 \| k_2 \| \dots \| k_{q'})$, 因此 q-composite 相对于基本随机密钥预分布模型来说对节点被捕获有很强的自恢复能力。在这一节, 我们将用我们的模型来分析当有节点被捕获时, 我们的模型如何加强 q-composite 密钥预分配模型的网络适应度。首先, 我们设前提条件为被捕获的节点数目为 x , 每个节点的密钥环为 m 。任意节点未被捕获的概率为 $(1 - \frac{m}{|S|})^x$, 因此网络通信信息被捕获率为 $1 - (1 - \frac{m}{|S|})^x$ 。对于任意两个节点的通信链路有 i 个共享密钥, 则链路节点被捕获率的方程为 $(1 - (1 - \frac{m}{|S|})^x)^i$ 。建立一条安全链路的概率为 $p = p(q) + p(q+1) + \dots + p(m)$ 。因此, 在有 x 个节点被捕获的情况下, 正常节点通信信息可被捕获的概率如式(15)所示[7]

$$f_q = \sum_{i=q}^m ((1 - (1 - \frac{m}{|S|})^x)^i \times \frac{p(i)}{p}) \quad (15)$$

要想安全网络中任意两点之间的安全连通率超过 q 的概率达到理想概率值 p (预先设定), 就必须缩小整个密钥池的大小, 增加节点间共享密钥的交叠度。但密钥池太大会使故

人捕获少数几个节点就能获得很大的密钥空间，因此寻找一个最佳的密钥池大小是 q-composite 的实施关键。

q-composite 随机密钥预分布模型密钥池的大小可以通过下面的方法获得[9]。假设网络的连通概率为 c ，每个节点邻居节点数的期望值为 n' ，根据式子(1)(2)可以得到任何给定的节点的连通度期望值 d 和节点间的连通率 p 。设 m 为每个节点存放密钥环的大小，要找到一个最大的密钥池 S ，使得从 S 的任意两次 m 个密钥采样，其相同密钥的个数超过 q 的概率大于 p 。设任意两个节点之间共享密钥个数为 i 的概率为 $p(i)$ ，则任意节点从 $|S|$ 个密钥中选取 m 个密钥的方法有 $C(|S|, m)$ 种，两个节点分别选取 m 个密钥的方法数为 $C^2(|S|, m)$ 个。假设两个节点之间有 i 个共同的密钥，则有 $C(|S|, i)$ 种方法选出相同的密钥，另外 $2(m-i)$ 个不同的密钥从剩下的 $|S|-i$ 个密钥中获取，方法数为 $C(|S|-i, 2(m-i))$ 。于是有[7]：

$$p(i) = \frac{C(|S|, i)C(|S|-i, 2(m-i))C(2(m-i), (m-i))}{C^2(|S|, m)} \quad (16)$$

为了便于计算，我们用 Stirling 约化公式^[11] $n! \approx \sqrt{2\pi n} n^{n+\frac{1}{2}} e^{-n}$ 化简公式(16)得：

$$p(i) = \frac{i!}{m^2} \times e^{2i} \times \left(\frac{|S|}{|S|-m}\right)^{|S|} \times \left(\frac{|S|+i-2m}{|S|-m}\right)^{|S|} \quad (17)$$

用 p_c 表示任何两个节点之间存在至少 q 个共享密钥的概率，则有[7]

$$p_c = 1 - (p(0) + p(1) + p(2) + \dots + p(q-1)) \quad (18)$$

通过式(18)我们可以根据 $p_c \geq p$ 计算出最大的密钥池尺寸 $|S|$ 。

接下来我们进行适应度抗攻击分析，文献[7]分析了网络规模为 n 的网络，在有 x 个节点被捕获的情况下，我们想要找到一个 m ，使得在捕获 x 个节点以后，网络的适应度 R_q 大于一个理想的值 R_{q0} ，该值的大小和环境和应用所需要的安全程度有关，即

$$R_q = 1 - \sum_{i=q}^m \left((1 - (1 - \frac{mq}{|S|})^x)^i \times \frac{p(i)}{p} \right) \geq R_{q0} \quad (19)$$

我们令 $1 - (1 - \frac{mq}{|S|})^x = t$

则式 (19) 化简为

$$R_q = 1 - \frac{t^q}{p} [p(q) + tp(q+1) + t^2 p(q+2) + \dots + t^{m-q} p(m)] \geq R_{q0} \quad (20)$$

由于 $t \leq 1$

我们对(21)进行放大得：

$$1 - \frac{t^q}{p} [p(q) + p(q+1) + p(q+2) + \dots + p(m)] \geq R_{q0} \quad (21)$$

最后化简得：

$$R_{qq} = 1 - (1 - (1 - \frac{mq}{|S|})^x)^q \geq R_{q0} \quad (22)$$

文献[7]提出提高 q 值可以提高系统的抵抗力，攻击网络的攻击难度和共享密钥个数 q 呈指数关系，与式(22)相符。

则

$$m_q \leq |S| (1 - (1 - (1 - R_{q0})^{\frac{1}{q}})^{\frac{1}{x}}) \quad (23)$$

因此，我们可以得到 m_q 的最大值

$$m_{q \max} = |S| (1 - (1 - (1 - R_{q0})^{\frac{1}{q}})^{\frac{1}{x}}) \quad (24)$$

为了保证网络的适应度控制在理想范围， m_q 的值不能超过 $m_{q \max}$ 。和前面讨论的相似，设 m_0 为密钥初始化阶段密钥环的初始大小； p 为节点间的连通率，和 m_q 的取值正相关；同时密钥池的大小 $|S|$ 是给定的。例如，如果设式 (23) 中 $p = 0.33$ ， $m_q = 200$ ， $R_{q0} = 0.8$ 。如果 m_q 小于 $m_{q \max}$ 将确保网络的适应度大于或等于 R_{q0} 。但是，如果密钥环的大小 m_q 较小，会导致网络连通率 p 降低。为了很好的解决适应度和节点间连通率之间权衡的问题，我们提出了类似于基本框架改进的方案，先设定密钥环的大小 m_q' 为变量，由于 m_q' 的变化，引起适应度 R_q 和两节点间连通率 p 发生变化。

$$R_q' = 1 - (1 - (1 - \frac{m_q'}{|S|})^x)^q \quad (25)$$

$$p_q'(i) = \frac{i!}{m^2} \times e^{2i} \times (\frac{|S|}{|S| - m'})^{|S|} \times (\frac{|S| + i - 2m'}{|S| - m'})^{|S|} \quad (26)$$

$$p_q' = 1 - (p_q'(0) + p_q'(1) + p_q'(2) + \dots + p_q'(q-1)) \quad (27)$$

如果 ($m_{q \max} < m_0$)，则

$$m_q' = m_{q \max} \quad (28)$$

式 (26) 和 (27) 中，如果 ($p_q' \leq p_{\min}$)，使

$$m_q' = m_{q \min} \quad (29)$$

其中 p_{\min} 为节点间连通率的最小阈值，该阈值的作用是为保证全网的连通率可以正常通信^[12]。

如果 ($m_{q \max} \geq m_0$)，则使得

$$m_q' = m_0 \quad (30)$$

在本文对改进的 q -composite 模型仿真中，我们给定 $p_{\min} = 0.6p$ ， $m_0 = 200$ ， $R_{q0} = 0.8$ ， $q = 1, 2, 3$ ， $p = 0.33$ 。

3 仿真结果和分析

我们把改进的结果同基本模型和 q -composite 模型对比来进行评估。图 2 展示了我们改进的模型同基本框架和 q -composite 中 $q = 1$ 模型的对比结果。在该图中的 4 个模型中 $|S| = 100000$ ， $m = 200$ ， $p = 0.33$ ， $N = 1000$ ， $n' = 100$ ， $R_0 = 0.8$ ， $R_{q0} = 0.8$ 。从图中我们可以看到，随着被捕获节点数增多网络适应度有持续的提高，但是整体的下降的趋势是不可避免的。由于 q -composite 要求以实际的共享密钥数 q' ($q' > q$) 通过 hash 函数来产生共享密钥，因此适应性同基本框架相比有所提高，但比我们改进后的基本模型仍有一定差距。 $R = 0.8$ 在图中是个重要的分界点，该点是我们保证该环境该应用理想安全度的临界点。从图中我们不难发现当改进

$q=1$ 的模型被捕获的节点数为 103 个的时候,已经无法满足适应度 $R \geq 0.8$ 的要求,因此我们可以理解该网已经不再安全,或者说达不到我们的安全要求。而改进的基本框架,可以容忍 238 个节点被捕获。

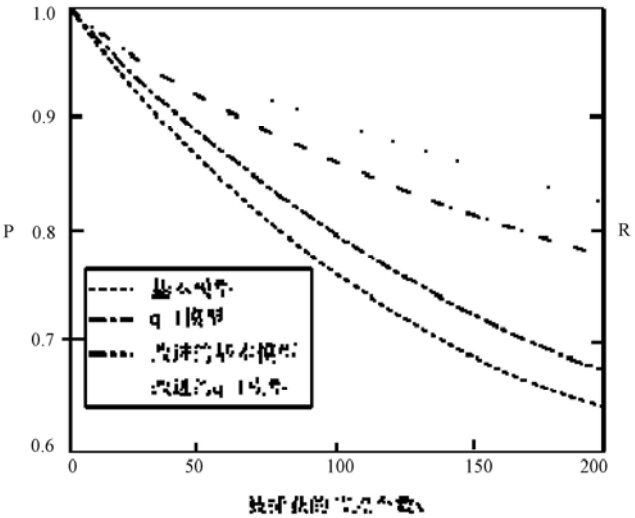


图 2 改进前后的基本模型和 q -composite 模型中,随着被捕获节点数目的增加,网络适应度的变化

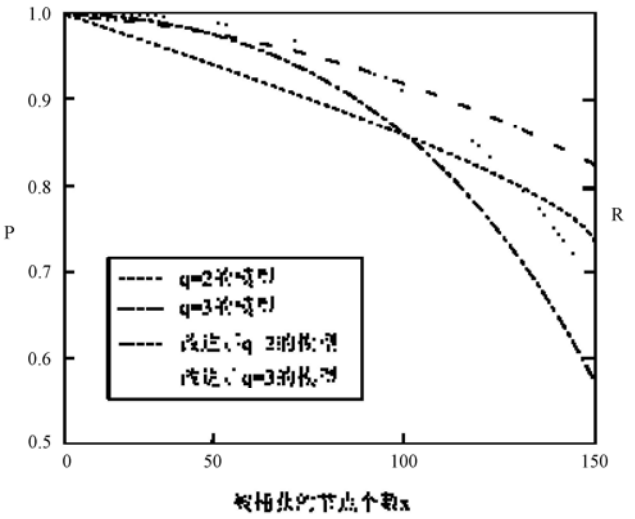


图 3 q -composite 模型中改进前后 $q=2$ 和 $q=3$ 模型,随着被捕获节点数目的增加,网络适应度的变化。

同样图 3 的条件为 $m=200, p=0.33, N=10000, n'=100, R_{q_0}=0.8$ 。通过仿真我们不难发现在适应度方面我们改进的结果比 $q=2$ 和 $q=3$ 模型都有一些提高,但由于考虑到传感器节点的低计算能力,我们公式 (19) 进行了简化,便于传感器节点的计算,因此改进的效果不如改进的基本框架那么明显。需要着重说明的一点改进前后 $q=2$ 和 $q=3$ 模型都存在交点,原因是通过给定的条件可以计算出 $q=2$ 和 $q=3$ 的密钥池大小, $|S|_{q=2}=67736$, $|S|_{q=3}=7613$, 因此,当被捕获节点增大到一定程度后, $q=3$ 被泄露的密钥数占密钥池的比例远远大于 $q=2$ 被泄露密钥数占密钥池的比例,因此 q 的值越大,随着节点被捕获数增大,其

安全性就越低。我们改进的模型仍然符合这个规律。同样值得注意的是在满足 $R_q \geq 0.8$ 的安全前提下改进后的两个模型比原模型容忍被捕获节点的数有所增加。说明模型的抗攻击能力有所增强。

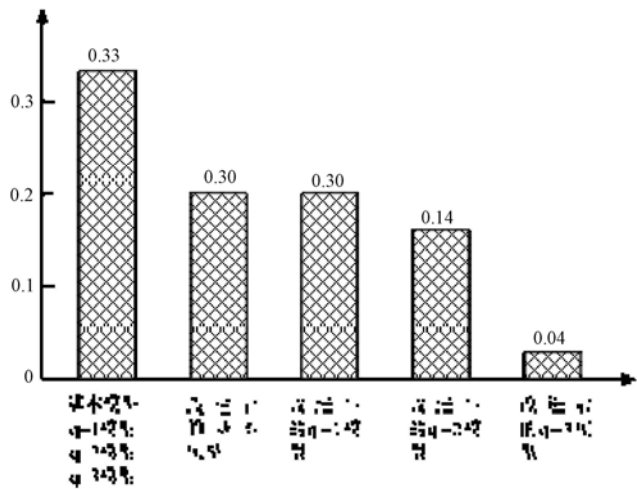


图 4 本文仿真结果所涉及到的任意两点连通率图

但是适应度 R 的提高是以牺牲节点间连通率为代价的。因此其改进协议能不能被接受，需要评估其连通率降低的程度。从图 4 中我们不难发现在邻居节点为 100 个的条件下，改进的基本模型和 $q=1$ 模型的任意两点间连通率为 0.20，带入公式(1),(2)得全网连通率 $c=0.99992$ ，能够保证 10000 个节点的完全连通。而 $q=2$ 的任意两点间的连通率为 0.16 带入式(1),(2)得 $c=0.9654$ ，而 $q=3$ 的全网连通率 $c=0.371$ 说明网络的通信空洞非常多，已经无法接受。对于邻居节点期望值不大的情况下，对于该模型如何在 q -composite 使用的方法有两个：1，通过式(1),(2),(16),(18)计算出一个适合该网络的 $p_{q\min}$ ，但对于一般规模 ($N < 10000$) 的传感器网络，两点间连通率可降低的程度非常有限。2，可以调节发射功率扩大通信半径来搜索能够通信的节点，当路径密钥形成阶段完成后，把发射功率调节到初始状态。总体来说，如果邻居节点数的期望值非常高，例如 $n'=2000$ ，该模型的效果就非常明显，通过计算可以得到 $q=2, q=3$ 的全网连通率都为 0.99999。可见该模型对邻居节点数期望多的网络支持性相对较好。

此外图 5 还展示了当 x 个节点数被捕获的前提下，正常节点被捕获的概率图。可见改进的基本模型的捕获率与原始的几个模型的捕获率相比，随着被捕获节点数的增长，其结果是最好的。定量来分析当邻居节点的期望值 $n'=100$ 的时候，我们改进的基本模型与原模型的节点连通率 p 相比损失了 38%，而全网的连通率 c 仅仅减少了 0.08%，改进的 $q=1$ 与基本模型全网连通率损失程度同为 0.08%，结合图 3 所得的数据我们分析可得 $q=2$ 损失全网连通率为 3.5%， $q=3$ 损失的全网连通率高达 63%。但是，如果邻居节点的期望值 $n'=1000$ 以上后，该改进模型相对于原模型损失的全网连通率都小于十万分之一。

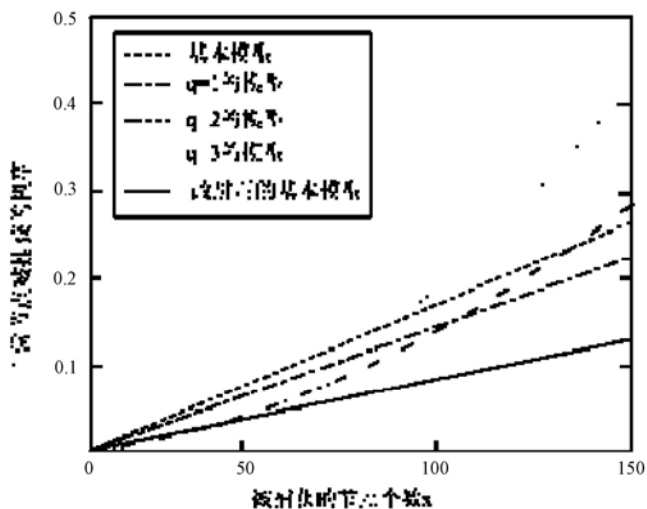


图 5 随着被捕获节点数与正常节点被捕获概率之间的关系图

4 结论

本文针对无线传感器网络提出一种可变的随机密钥预分配模型，通过引入适应度这个概念来利用抗攻击模型，对已有节点被捕获的网络进行优化。从安全性来考虑，该模型不仅仅提高了部分节点被捕获网络的生命力，还普遍适用于随机密钥预分配模型，不仅仅局限于基本模型，对于 q -composite 也有一定的改进。仿真结果证实，该抗攻击模型对于满足大规模条件的传感器网络尤其适用。

参 考 文 献

- [1] Carman D W, Kruus P S and Matt B J, Constraints and Approaches for Distributed Sensor Network Security[M], NAI Labs Technical Report No.00-010, Sept, 2000
- [2] Akyildiz I.F, Su W, Sankarasubramaniam Y, A Survey on Sensor Networks[J], IEEE Communications Magazine, 40(8), 102~114, 2002
- [3] 苏忠, 林闯, 封富君 无线传感器网络密钥管理的方案和协议[J], 软件学报, 18 (5), 1218~1231, 2007,
- [4] Du W, Deng J, Han Y S, *et al*, A Key Management Scheme foe Wireless Sensor Networks Using Deployment Knowledge[C], IEEE INFOCOM, 2000
- [5] Perrig A, Stankovic J, Wagner D, Security in wireless sensor networks [J], Communications of the ACM, 47(6), 53~57, 2004
- [6] Eschenquer L and Gligor V D, A key-management scheme for distributed sensor networks[J], Proc. of the 9th ACM Conf on Computer and Communication Security, 41~47, Nov. 2002
- [7] Chan H, Perring A, Song D, Random key predistribution schemes for sensor network[C], IEEE Symp on Security, 197~213, May. 2003
- [8] Spencer J, The Strangle Logic of Random Graphs, Algorithms and Combinatorics [M] 22 Springer-Verlag, ISBN 3-540-41654-4, 2000
- [9] 孙利民, 李建中, 陈渝, 无线传感器网络[M], 北京, 清华大学出版社, 2005

- [10] Carl H T, James B and Richard H, Node Compromise in Sensor Networks: The Need for Secure Systems[C], Technical Report CU-CS-990-05, Dept of Comp. Sci., Univ. of Colorado at Boulder, 2005
- [11] Wikimedia Foundation, Inc, Stirling's approximation [EB/OL], http://en.wikipedia.org/wiki/Stirling's_approximation, 2009
- [12] Lu K J, Qian Y, Guizani M, *et al*, A Framework for a Distributed Key Management Scheme in Heterogeneous Wireless Sensor Networks[J], IEEE Transactions on Wireless Communications, 7(2), 520~526, 2008

作者简介

张啸（1983—），男，汉族，深圳大学信息工程学院 硕士研究生，研究方向：无线传感器网络安全技术
xiaozh@szu.edu.cn

王晖（1969—），男，汉族，深圳大学信息工程学院 教授，硕士生导师，研究方向为信号处理及无线通信。长期从事数字图象处理、智能医学信息处理和数字通信等方面的研究，主持并完成了多个国家及省、市自然科学基金项目，是国际国内多个核心期刊审稿人，并且在国际国内一流学术期刊与会议上发表论文二十余篇。wanghsz@szu.edu.cn

林晓辉（1975—），男，汉族，深圳大学信息工程学院 教授，研究生导师，研究方向为无线网络与分布式系统。长期从事移动计算与分布式系统，无线网络资源管理，无线网络功率控制与流量调度，路由算法设计与评估，优化理论与应用等方面的研究，主持多个国家及省、市自然科学基金项目、深港创新圈项目等，并在 IEEE 等国际一流学术期刊与会议上发表论文二十余篇。

基于加权二乘向量机的图像增强算法

彭代强 林幼权

(南京电子技术研究所, 南京 210013)

摘要: 提出了一种基于二乘向量机的自适应图像增强方法, 该方法通过将图像空间映射到一个高维核空间, 进而在高维核空间中描述图像灰度信息, 同时针对图像的局部灰度特征, 在二乘向量机中构建一个自适应惩罚因子来控制误差函数, 以提高图像降噪性能。仿真结果表明, 优化后的方法能够有效的增强图像信息, 较好地保护边缘细节。

关键词: 统计学习理论; 最小二乘向量机; 图像增强; 边缘保持

Image Enhancement Using Weighted least squares support vector machines

Peng Dai qiang Lin You quan

(Nanjing Research Institute of Electronics Technology, Nanjing 210013, China)

Abstract: The paper presents an adaptive new image enhancement approach based on least square support vector machine (LS-SVM). The approach is to map the data into higher dimension space in which the information of input image can be represented. According to the local gray feature of noisy image, we construct adaptive weighted factors in LS-SVM to control the error function. It improves the performance of noise reduction. Experimental results demonstrate that the proposed technique is efficiency of image enhancement and can better preserve sharp edge.

Keywords: Statistical Learning Theory(STL); Least Square Support Vector Machines (LS-SVM); Image enhancement; Edge preservation

图像在其采集、传输过程中会不可避免地受到噪声的污染, 噪声对图像的干扰严重地影响了图像的质量, 给图像观测、特征信息提取和分析等图像处理过程带来了极大的困难。因此, 长期以来, 消除和降低图像中的噪声一直是图像处理领域感兴趣的课题之一。图像降噪的目的是在尽可能多地保留图像特征的前提下去除噪声污染, 现有的许多方法多采用平滑滤波的机理, 如均值滤波, 高斯滤波等, 这些方法在降低图像噪声的同时, 会不可避免的对图像的细节及边缘信号造成损坏, 因而怎样在图像降噪和图像细节信息保留之间采取有效的折衷是图像降噪的关键。

支持向量机是 Vapnik^[1]等人根据统计学习理论提出的一种学习方法, 近年来在模式识别、回归分析和特征提取等方面得到了很多应用。支持向量机方法根据 Vapnik 的结构风险最小化原则, 尽量提高学习机的泛化能力, 即是由有限训练样本得到的决策规则对独立的测试集仍能够得到小的误差, 是兼顾到经验风险和置信范围的一种折衷的思想, 为解决不等式约束的

二次优化问题，Suykenst^[2]在 SVM 的优化函数中引入误差二次项，并将 SVM 中的不等式约束条件改为等式约束，提出了一种以二次等式约束条件为基础的改进型向量机即最小二乘向量机。本文利用 LS-SVM 的回归原理，提出了一种自适应加权 LS-SVM 图像降噪方法，该加权因子随着图像的局部统计结构而变化，因此可以有效的解决降噪后图像所含噪声及降噪后图像细节保留程度之间的矛盾，试验结果表明该方法不仅能够有效地去除高斯噪声，而且能够很好的保留图像的细节纹理。

1 最小二乘向量机

由于 LS—SVM 通过误差二次项的引入将经典 SVM 优化函数的不等式约束改成了只有等式约束，这样 LS—SVM 的求解问题从标准 SVM 的二次函数寻优问题转换为线性方程求解问题，大大地简化了问题的复杂性。对于一组给定的训练样本 $\{(x_i, y_i)\}_{i=1}^N$ ，其中 N 为样本数， $x_i \in R^d$ 为输入数据， $y_i \in R$ 为输出数据，经修改后的 LS-SVM 目标函数为：

$$\begin{aligned} \min \quad & J = \frac{\|\omega\|^2}{2} + \frac{\gamma}{2} \sum_{i=1}^N e_i^2 \\ \text{s.t.} \quad & \\ & y_i = \sum_{i=1}^N \omega_i * \varphi(x_i) + b \end{aligned} \tag{1}$$

其中 $\varphi(\cdot)$ 为非线性变换函数，它将输入空间非线性变换到高维空间， e_i 为松弛项， γ 为常数，其为错误惩罚因子。相应的拉格朗日函数为：

$$L = \frac{\|\omega\|^2}{2} + \frac{\gamma}{2} \sum_{i=1}^N e_i^2 - \sum_{i=1}^N \alpha_i \{ \omega \varphi(x_i) + b - y_i + e_i \} \tag{2}$$

式中 α 为拉格朗日乘子，对于上式方程，根据 KKT 条件消去 ω 和 e 得到如下线性方程

$$\begin{bmatrix} 0 & \bar{1}^T \\ \bar{1} & K + \gamma^{-1}I \end{bmatrix} \begin{bmatrix} b \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ Y \end{bmatrix} \tag{3}$$

其中 $Y = [y_1; \dots; y_N]$, $\bar{1}^T = [1, \dots, 1]$, $\alpha = [\alpha_1; \dots; \alpha_N]$ ， K 为核函数 $K(x, x_i) = \varphi(x) \varphi(x_i)$ ， $i = 1, \dots, N$ ，这样基于支持向量的函数估计形式为

$$f(x) = \sum_{i=1}^N \alpha_i K(x, x_i) + b \tag{4}$$

其中 α 和 b 是等式(3)的解，具体形式取决于错误惩罚因子 γ 及核函数的类型。

2 基于最小二乘向量机的图像增强算法

假定 Y 为噪声图像， S 为无噪声干扰图像，在图像处理中，通常我们并不是直接用整幅噪声图像 Y 进行处理，而是通过图像像素邻域进行处理，典型的图像像素邻域的大小为 3×3 到 9×9 不等。很明显，在一个像素的邻域中，LS—SVM 的输入矢量由像素坐标定义，输出目标则为像素的灰度值，定义 W 为可移动的像素领域窗口， $W = R \times C$ ，其中 R 和 C 为图像中

某一像素对称邻域内的坐标序数集。以 RBF 为核函数的 LS-SVM 图像灰度函数可表示为^[3]:

$$f(x, y) = \sum_{i=1}^N \alpha_i \exp \left\{ -(|x - x_k|^2 + |y - y_k|^2) / \sigma^2 \right\} + b \quad (5)$$

式中 b 为窗口 W 内的局部均值, α 为窗口 W 内像素点对应的拉格朗日乘子, (r, c) 为窗口 W 内像素点坐标定义。随着窗口的移动, 我们以窗口中心点 $f(0,0)$ 作为滤波器输出。文献^{[4]-[7]}运用支持向量机对图像降噪进行了研究, 但这些传统的 SVM 降噪算法中错误惩罚因子 γ 为一常数, 从等式 (1) 可看出 γ 控制着图像平滑及边缘保留, 当 γ 较大时, LS-SVM 目标函数对近似误差惩罚较大, 因此图像边缘细节保持较好而噪声抑制效果较差; 当 γ 较小时, LS-SVM 目标函数对近似误差惩罚较小, 因此图像噪声抑制较为理想而边缘细节信息被平滑, 实际上在不同的图像信息结构中, 我们需要对图像边缘及细节信息进行保持, 而对均匀区域内的噪声尽可能平滑, 因此传统的 SVM 降噪算法存在着保护信号局部特征与抑制噪声之间的矛盾。为提高 LS-SVM 算法的降噪性能, 针对图像灰度特征的变化, 建立自适应错误惩罚因子以控制图像平滑及边缘保持性能, 可得到如下的目标函数:

$$\min J = \frac{\|\omega\|^2}{2} + \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N v_{i,j} e_{i,j}^2 \quad (6)$$

满足约束:

$$y_j = f(x, \omega) = \sum_{i=1}^N \omega_i \times \phi(x_i, x_j) + b + e_{i,j} \quad j=1, \dots, N \quad (7)$$

由文献^[8]所述, 其解可归结为以下线性方程组

$$\begin{bmatrix} 0 & \bar{1}^T \\ \bar{1} & \Omega + V_{\text{weighted}} \end{bmatrix} \begin{bmatrix} b \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ Y \end{bmatrix} \quad (8)$$

其中 $\Omega_{k,l} = K(x_k, x_l)$, $1/v_{i,j}$ 为矩阵 V_{weighted} 的第 (i, j) 个元素。事实上等式 (8) 给出了拉格朗日乘子 α 的解

$$\alpha = (\Omega + V_{\text{weighted}})^{-1} (Y - \bar{1}b) \quad (9)$$

于是基于最小二乘支持向量机的图像降噪算法求解就转化为如何选取自适应错误惩罚因子, 该惩罚因子不仅与像素领域内的灰度特性有关, 由于支持向量机使用非线性变换函数 $\phi(\cdot)$ 将输入空间映射到高维空间, 因此其还与核函数有关, 我们定义

$$V_{\text{weighted}} = T\Omega \quad (10)$$

其中 T 为与像素领域内的灰度特性相关的矩阵。从等式 (9) 我们注意到噪声图像 Y 加权后的拉格朗日乘子 α 与未受噪声干扰的图像 S 在没有加权情况下的拉格朗日乘子 α 相同, 于是

$$\alpha = (\Omega + V_{\text{weighted}})^{-1} (Y - \bar{1}b) = (\Omega)^{-1} (S - \bar{1}b) \quad (11)$$

由上式可得:

$$Y - S = \Gamma(S - \bar{1}b) \quad (12)$$

考虑噪声为零均值高斯白噪声及信号与噪声独立, 对上式求取滑动窗口 W 内的平方和, 可得:

$$T = \frac{\sigma_n}{\sqrt{\sigma^2 - \sigma_n^2}} \quad \text{if } \sigma^2 > \sigma_n^2 \quad (13)$$

上式中 σ_n^2 为噪声方差, σ^2 为滑动窗口 W 内噪声图像的局部方差。可以看出, T 随图像的局部方差而变化, 当图像的局部方差大于噪声方差时, 表明滑动窗口在图像的边缘区域, 我们需要增大近似误差的影响以保留图像的边缘特称, 也就是减小 T 值; 当图像的局部方差小于噪声方差时, 表明滑动窗口有可能在图像的均匀区域, 我们需要减少近似误差的影响以平滑图像, 此时我们选取一个较大的 T 值。由以上分析可得 T 为:

$$T_i = \begin{cases} \frac{\sigma_n}{\sqrt{\sigma_i^2 - \sigma_n^2}} & \text{if } \sigma_i^2 > \sigma_n^2 \\ 20 & \text{if } \sigma_i^2 \leq \sigma_n^2 \end{cases} \tag{14}$$

- 根据以上分析, 基于最小支持向量机的图像降噪算法归纳如下:
- (1) 移动窗口 W 统计窗口内的图像灰度方差作为窗口中心像素点方差, 计算整个图像所有像素点方差, 根据等式 (14) 计算图像中每个像素点对应的 T 值。
 - (2) 选取 LS-SVM 核函数类型, 根据 T 值及核函数, 由等式 (10) 确定窗口 W 内的像素点对应的自适应错误惩罚因子 V_{weighted} 。
 - (3) 利用窗口内噪声图像像素点作为训练点, 由等式 (9) 计算窗口内图像对应的拉格朗日乘子 α
 - (4) 利用 LS-SVM 回归函数式 (5) 对图像处理区域内的像素点进行灰度估计, 即图像降噪, 将窗口中心点的估计输出作为图像滤波输出。
 - (5) 移动窗口选取下一块图像处理区域, 返回第 (2) 步直到全部图像处理完毕。

3 实验结果

为验证所提出的基于 LS-SVM 图像降噪算法的有效性, 我们采用 256×256 的 Lena 和 Toys 图像作为测试图像, 滑动窗口 W 大小选取为 3×3 , 即像素点坐标定义为 $R = \{-1, 0, 1\}$, $C = \{-1, 0, 1\}$, 在支持向量机模型中选取高斯径向基函数 (RBF) 作为核函数。在实验中采用滤波前后图像的峰值信噪比 (PSNR) 作为降噪评价标准, 峰值信噪比 PSNR 定义为

$$\text{PSNR} = 10 \log \left(\frac{\sum 255^2}{\sum [S(i, j) - f(i, j)]^2} \right) \text{dB} \tag{15}$$

为了进行比较, 对测试图像分别加入标准差为 10、15、20、25、30 的高斯白噪声, 同时也采用传统的 LS-SVM 滤波方法、中值滤波及小波软阈值方法进行降噪仿真, 降噪仿真对比结果见表 (1)。由表 (1) 可以看出, 在不同信噪比下, 本文提出的算法得到的信噪比变化较大, 相对于传统的支持向量机去噪算法, 图像信噪比提高了 0.58dB 到 1.9dB, 相对于其它降噪算法, 本文提出的算法在噪声较小的情况下, 降噪效果较为理想, 而小波软阈值方法在噪声较小时, 图像信息损失较大。

表 1 四种算法下 PSNR 的比较

噪声图像	传统 LS-SVM	中值滤波	小波软阈值	本文算法
28.15	29.96	29.14	27.90	31.86
24.66	28.94	27.98	26.26	30.10
22.14	26.51	26.76	25.32	28.95
20.23	26.03	25.63	24.18	26.91
18.64	25.18	24.53	23.83	25.76

图（1）示出了峰值信噪比 PSNR 为 22.14 的噪声图像及降噪后的图像结果。从图中不难看出用中值滤波降噪时，噪声没有得到有效抑制，传统的支持向量滤波方法边缘保持不是较好，本文提出的最小支持向量机图像降噪方法能有效的滤除噪声，同时具有更好的细节保持特性，图像具有更好的清晰度。



图 1 不同算法的图像增强性能比较

图（2）为 Toys 图像的滤波效果，为对比本文算法和传统 LS-SVM 算法的滤波效果，针对图中方框图像区域的滤波结果，进行了局部放大处理，图 2（g）为本文提出的自适应惩罚因子，该因子随着图像的局部细节信息发生变化，可以看出本文算法有效的保留了图像的细节信息。

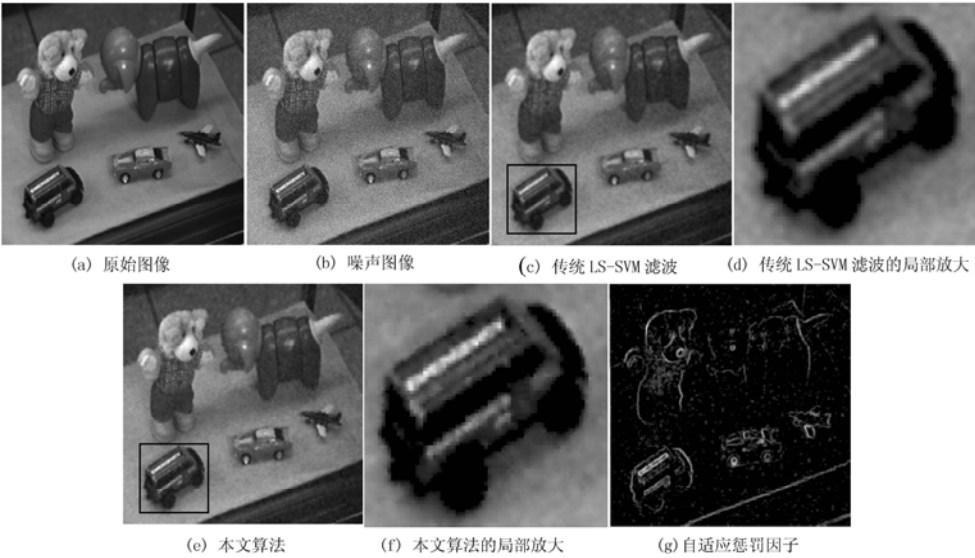


图 2 Toys 图像处理结果

4 结束语

传统的支持向量机降噪算法对整幅图像使用同一错误惩罚因子控制误差函数,当图像局部特征发生变化时,存在着保护信号局部特征与抑制噪声之间的矛盾。本文提出的最小二乘向量机图像增强方法根据图像灰度特征的变化,自适应调整错误惩罚因子来控制回归误差,从而达到最佳滤波效果。实验结果表明,该算法具有较好的噪声抑制及保持边缘能力。

参 考 文 献

- [1] Vapnik V. An Overview of Statistical Learning Theory[J]. IEEE Transactions Neural Networks,1999,10(5): 988 - 999.
- [2] Suykens J A K, Vandewalle J. Least squares support vector machine classifiers[J]. Neural Processing Letters, 1999, 9(3):293-300
- [3] 郑胜,田金文,柳健. 二乘向量机在图像插值技术中的应用研究明. 计算机工程与应用[J], 2004, 17(1): 82-84
- [4] Dick K T C, Tong L. Image approximation and smoothing by support vector regression[J]. IEEE on Neural Networks Conference Proceeding, 2001:2427-2432
- [5] 周广辉 黄炜. 基于 LS-SVM 的图像去噪方法 电脑应用技术[J], 2006 65:15-18
- [6] Han Liu, Yong Guo, Gang Zheng. Image denoising based on least square support vector machines [C]. IEEE on Intelligent Control and Automation Conference Proceeding, 2006:4181-4184
- [7] Chunyu Chen, Xiaohui Qi, Maoliu Lin. Research and application of noise suppression based on support vector machine[C]. International Symposium on Communications and Information Technologies, 2005: 346-349
- [8] Suykens J A K, DeBrabanter J, Lukas L, Vandewalle J. Weighted least squares support vector machines: robustness and sparse approximation[J]. Neurocomputing, 2002(48): 85-105

作者简介:

彭代强(1975—),男,四川苍溪,博士研究生,南京电子技术研究所,主要研究方向图像处理;

林幼权(1966—),男,浙江余姚,博士研究生,研究员,南京电子技术研究所,主要研究方向雷达系统设计

基于置信度传播的MSTAR图像的分割算法

徐 戈 黄志强

(南京电子技术研究所, 江苏 南京 210039)

摘 要: 由于 SAR 图像具有斑点噪声, 基于强度的分割方法往往失效。针对该问题, 引入马尔可夫随机场框架, 设计了一种能量函数, 该能量函数将图像数据的一致性与空间分布约束统一起来, 有效克服 Speckle 影响; 在后续的函数优化阶段, 采用置信度传播算法来快速求解。一系列结果证明, 此方法分割精度高, 计算速度快, 有效解决 MSTAR 图像分割问题。

关键词: 马尔可夫随机场; 置信度传播; 分割; SAR 图像

A New Segmentation Algorithm of MSTAR Imagery based on Belief Propagation

Xu Ge Huang Zhi qiang

(Nanjing Research Institute of Electronic Technology, Nanjing 210039, China)

Abstract: Due to Speckle noise, purely intensity-based segmentation technique for SAR imagery failed. We introduced Markov random fields framework, and proposed a new energy function, where the energy function had two terms: one term penalized solutions that are inconsistent with the observed data, whereas the other term enforced spatial coherence. In energy optimization, we applied belief propagation approach that substantially improved the running time. A series of preferable result demonstrate that our algorithm has two advantages: accurate and fast, and it is adapt to segmentation for MSTAR imagery.

Keywords: Markov Random Field; Belief Propagation; Segmentation; SAR Imagery

1 引言

合成孔径雷达(synthetic aperture radar, SAR)具有全天候、全天时和穿透能力强的特点, 在军事和民用领域均获得广泛的关注。在军事上, 该技术是一种获取地面信息的重要手段, 研究 SAR 图像中军事目标的自动目标识别(automation target recognition, ATR)技术, 对战场信息的描述、表示、理解和评估是十分重要的。在自动目标识别中, 分割技术是关键技术, 其实现水平直接决定 SAR 系统 ATR 的性能。

遗憾的是, 迄今为止, SAR 图像的分割问题仍然是未能圆满解决的关键问题。该问题难点主要归结为两个方面: 建模与优化。在建模方面, 基于马尔可夫随机场(Markov random field, MRF)框架的方法得到了广泛的应用。该框架可以将图像数据的一致性与空间分布约束统一起来[1]。在该框架下, 分割问题归结为 MRF 的最大后验概率的估计[2][3]。一系列研究表明,

在分割算法使用该框架可以有效克服 SAR 图像的斑点噪声影响。尽管 MRF 框架具有优雅性与有效性，但是其优化算法的效率直接限制了该框架的实用性。早期的优化算法，不论是迭代条件模式 (iterated conditional modes, ICM)^[2] 还是模拟退火(simulated annealing, SA)^[4] 都被证明要么无效，要么效率低下。针对这一问题，很多研究者展开了工作，研究了一系列的快速优化算法，如 graph cuts^[5] 和 belief propagation (BP)^[6]。

本文在前人研究的基础上，针对实际的运动和静止目标获取与识别(moving and stationary target acquisition and recognition, MSTAR)的高分辨率 SAR 图像，设计了适合 SAR 图像分割的能量函数，并采用 BP 完成能量优化。

本文其余章节安排如下，第 2 节讨论分割问题的 MRF 表达式，并设计适于 MSTAR 数据的分割能量函数，第 3 节简要介绍 BP 算法及其算法流程，第 4 部分给出具体的实验数据结果和分析，最后对全文进行总结。

2 分割模型

2.1 能量模型

本文将 MSTAR 图像分割问题定义对图像中的每个像素 p 进行标号，记为 f_p ，像素个数记为 n ，标号的个数记为 m ，能量函数 E 可以视为对 MRF 的后验分布取对数，该能量包括数据能量 E_d 和平滑能量 E_s 两项，记作：

$$E = E_d + \lambda E_s \quad (1)$$

其中，参数 λ 为平滑能量项调和系数；数据能量是每个像素数据代价 $d_p(f)$ 的和，记作：

$$E_d = \sum_p d_p(f_p) \quad (2)$$

在 MRF 框架中，数据能量为对噪声测度取负对数结果。本文针对的是 2 维栅格图像，像素 p 可以记为坐标的形式 $p(i, j)$ 。若采用 4 邻域系统，平滑项能量是空间中水平方向和垂直方向上最近邻域的平滑代价之和。最近邻域平滑项记为 $V_{pq}(f_p, f_q)$ ，其中 $p(i, j)$ ， $q(s, t)$ ，且 $|i-s|+|j-t|=1$ 。定义 N 为所有相邻像素对集合，平滑能量可以记为

$$E_s = \sum_{\{p,q\} \in N} V_{pq}(f_p, f_q) \quad (3)$$

在 MRF 框架中，平滑项能量描述的是先验信息取负对数的结果。本文所定义的平滑项代价与相邻标号的差值有关，具体如下式所示：

$$E_s = \sum_{\{p,q\} \in N} w_{pq} V(|f_p - f_q|) \quad (4)$$

其中，平滑能量项为空间权重 w_{pq} 与以标号差为变量的非增函数 $V(\cdot)$ 的乘积。

2.2 能量函数设计

MSTAR 图像数据由圣地亚(Sandia National Laboratory)国家实验室分别在 1995 年和 1996 年利用 X 波段、单视、HH 极化方式、0.3m×0.3m 分辨率，聚束 SAR 采集而得，其中有各种姿态角下的 10 种军用车辆的图像。图 1 给出了该数据集的部分图像，其尺寸为 128×128，

其中图像 1(a)中的目标为 BMP 装甲车, 图像 1(b)中的目标为 T72 坦克。

首先进行数据能量设计。从这两组原始图像可以看到, 图像是由背景杂波、阴影和目标三部分构成的。通过对这三部分进行取样, 分别得到三部分直方图, 如图 2 所示。

从图中可以看出背景杂波呈明显的重尾(heavy tail)分布, 进一步研究表明, 背景杂波服从有效视数为 1 的 Gamma 分布:

$$p_{\text{cluster}}(I_p, \mu) = \frac{L^L I_p^{L-1}}{\Gamma(L) \mu^L} \exp\left(-\frac{L I_p}{\mu}\right) \quad (5)$$

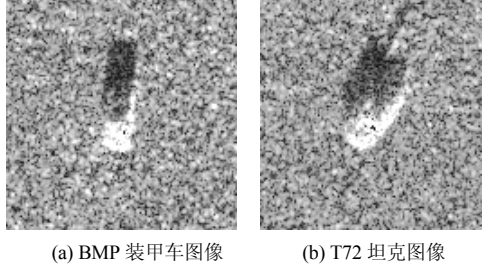


图 1 MSTAR 图像数据

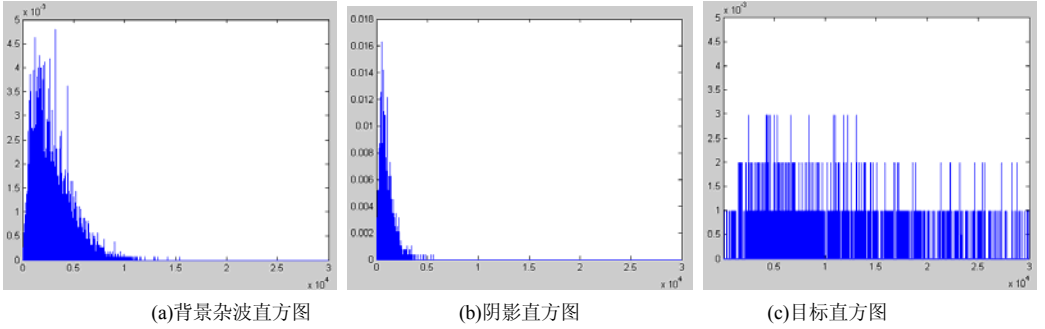


图 2 各部分直方图

其中 $L=1$, μ 为 Gamma 分布的参数, $\Gamma(\cdot)$ 为 Gamma 函数。

进一步分析, 阴影部分也有一定的重尾, 也可用有效视数为 1 的 Gamma 分布来近似:

$$p_{\text{shadow}}(I_p, \mu) \approx \frac{L^L I_p^{L-1}}{\Gamma(L) \mu^L} \exp\left(-\frac{L I_p}{\mu}\right) \quad (6)$$

目标部分的均值与方差较大, 本文采用高斯分布来近似:

$$p_{\text{target}}(I_p, \mu, \sigma) \approx \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(I_p - \mu)^2}{2\sigma^2}\right) \quad (7)$$

值得注意的是, 目标与阴影的分布均为近似分布, 不是其真实分布。分布上的近似主要是为了简化, 实践证明, 这样的简化是行之有效的。因而, 数据能量可以记作:

$$d_p(f_p) = -\log p(I_p) \quad (8)$$

本文采用 Potts 模型作为平滑能量。其数学表达式为:

$$V_{pq}(f_p, f_q) = w\{p, q\} \delta(f_p \neq f_q) \quad (9)$$

其中 $w\{p, q\}$ 是集团 $\{p, q\}$ 的加权系数, 该数是一个非负数。 $\delta(x)$ 定义如下:

$$\delta(x) = \begin{cases} 1 & x = 0 \\ 0 & \text{else} \end{cases}$$

其中惩罚系数与势团 $\{p, q\}$ 中的两像素的具体灰度值有关。如下设计 $w\{p, q\}$:

$$w_{p,q} = \exp(-\frac{\Delta I}{2\sigma^2}) \quad (10)$$

其中参数 σ 在这里被称为平滑项空间加权系数中的控制参数, ΔI 为像素间的灰度差。

总结公式(1)(8)(10), 得到本文分割能量函数:

$$E = \sum_p -\log p(I_p) + \lambda \sum_{\{p,q\} \in N} \exp(-\frac{|I_p - I_q|}{2\sigma^2}) \delta(f_p \neq f_q) \quad (11)$$

不难发现, 该能量函数的最小值求解是一个典型 NP 问题, 为了兼顾精度与效率, 本文采用置信度传播进行优化。

3 置信度传播(Belief Propagation, BP)算法

置信度传播算法有两种版本, 最大积(max-product)算法与和积(sum-product)算法。一般而言, 后者不是直接用于最小能量求解的, 该算法用于图中每个节点的边缘概率计算; 最大积多用于最大后验概率的求解。这里, 本文采用最大积算法近似求解式(11)的最小值。一般而言, 最大积算法定义在概率分布上, 通过取负对数, 最大积算法可以转化为最小和(min-sum)算法。

最大积 BP 算法是通过图像像素格点间的消息传递来完成的, 该方法是一个迭代、并行的算法。每个消息都是一个矢量, 其维数为该格点可能的标号, 在每次迭代中, 新消息按照如下方式计算:

$$m_{p \rightarrow q}^t(f_q) = \min_{f_p} \left(\lambda V_{p,q}(f_p, f_q) + d_p(f_p) + \sum_{s \in N(p) \setminus q} m_{s \rightarrow p}^{t-1}(f_p) \right) \quad (12)$$

这里 $s \in N(p) \setminus q$ 表示像素 p 的邻域但不包括像素 q , 经过了 T 次迭代, 每个节点的置信度矢量按照如下公式计算

$$b_q(f_q) = d_q(f_q) + \sum_{p \in N(q)} m_{p \rightarrow q}^T(f_q) \quad (13)$$

置信度矢量 $b_q(f_q)$ 中的最小值所对应的标号就是优化后的最优标号 f_q^* 。标准的消息传递算法的计算时间代价为 $O(nk^2T)$, 其中 n 为像素点数, k 为每个像素的可能标号数, T 为迭代次数。与经典的模拟退火不同, 该算法是一种确定性算法, 不需要设置繁琐的参数(在模拟退火算法中, 必须小心设置初始温度与退火策略), 这也是本文采用 BP 算法的原因之一。

4 参数设置与试验结果

观察分割能量函数, 其中涉及的参数有: 分割类别数 N , 数据能量项中的分布参数(具体指的是杂波区均值、阴影区均值, 目标区均值与方差), 平滑能量项中的参数 σ 与调和系数 λ 。对于MSTAR数据而言, 分割类别数 $N = 3$, 平滑能量项中的参数 σ 与调和系数 λ 可以通过实验确定出合适的参数。值得注意的是, 由于SAR图像成像的特殊性, 无论是阴影, 地面杂波,

还是目标，不同图像中的分布参数均是不同的(见表 1)，因此目标的分布参数既不能事先确定也不能通过试验固定。这里，本文采用人机交互的方法，即手工标定部份目标来获得目标的分布参数，如 图 3 所示，其中红色、绿色和黄色表示人工标定的区域，分别代表目标区域，地面杂波区域和阴影区域。

表 1 不同区域类别的分布参数

类别 \ 编号	hb03333	hb03334	hb03335	hb03337	hb03338	hb03343
目标区域均值	15717	13609	10483	15483	17175	12927
阴影区域均值	1200.4	891.54	880.03	953.37	1047.7	715.82
杂波区域均值	3218.1	2790.8	2164.3	2939.7	3298.5	2010.4

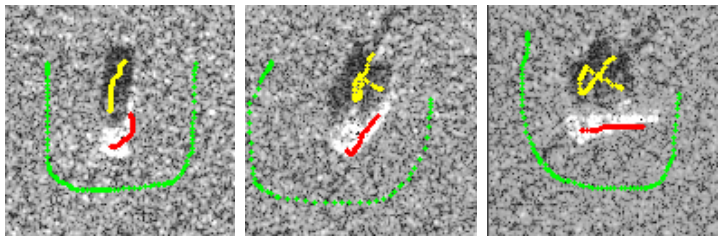


图 3 交互取样示意图

本文选取了两类目标不同的图像进行分割，原始图像如图 5 中的目标为 T72 坦克，具体参数设置见表 2。通过比对可见，本文提出的方法克服了 Speckle 的影响，将目标区域和阴影区域完整地从杂波中分割出来，为后续的目标识别提供了可靠的信息。更为重要的是，该算法是一种准实时算法，每幅图像的计算时间仅为 0.1 秒左右(CPU 为 Intel Pentium2.0GHz，内存 1G)。

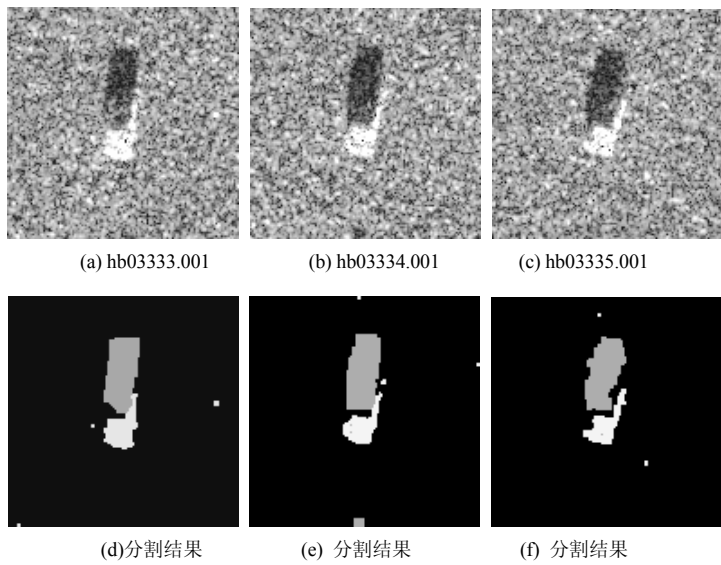


图 4 MSTAR 图像与相应的分割结果(BMP 装甲车)

表 2 参数设置

类别参数	$N = 3$
平滑项空间加权系数中的控制参数	$\sigma^2 = 5e^5$
平滑能量项调和系数	$\lambda = 1$

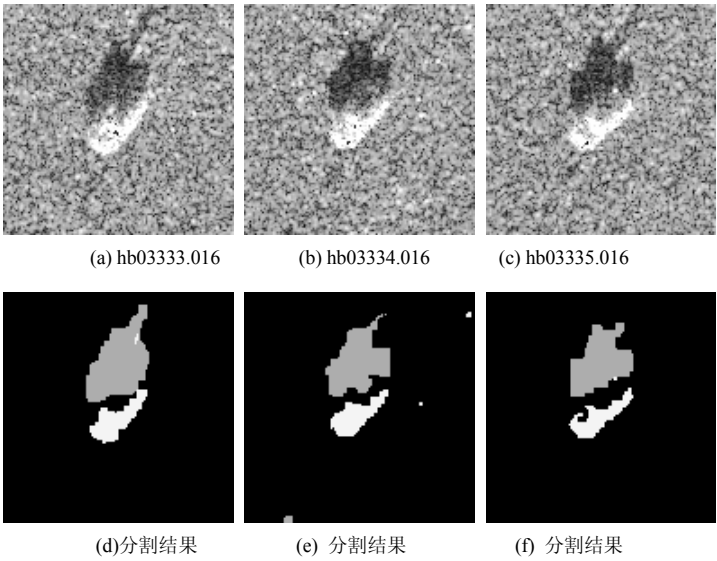


图 5 MSTAR 图像与相应的分割结果(T72 坦克)

5 结论

本文在 MRF 框架下，设计了一种适合 SAR 图像目标分割的能量函数；在能量优化阶段，采用置信度传播算法快速求解，从而有效地解决了 MSTAR 图像中的目标分割问题。与传统的分割算法相比，本文的算法具有分割精度高、计算复杂度低、参数设置简单的特点，这些特点使得该方法完全适用于实际的 SAR ATR 系统。

参 考 文 献

[1] Richard Szeliski, Ramin Zabih, Daniel Scharstein, Olga Veksler, Vladimir Kolmogorov, , Aseem Agarwala, Marshall Tappen, and Carsten Rother, “A Comparative Study of Energy Minimization Methods for Markov Random Fields with Smoothness-Based Priors”, IEEE Trans. Pattern Analysis and Machine Intelligence[J], vol. 30,no. 6, pp 1068-1080,June 2008

[2] J. Besag, “On the Statistical Analysis of Dirty Pictures (with discussion),” J. Royal Statistical Soc., Series B, vol. 48, no. 3, pp. 259-302, 1986.

[3] S. Geman and D. Geman, “Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images,” IEEE Trans. Pattern Analysis and Machine Intelligence[J], vol. 6, pp. 721-741, 1984.

[4] S. Barnard, “Stochastic Stereo Matching over Scale,” Int’l J.Computer Vision, vol. 3, no. 1, pp. 17-32, 1989.

[5] Y. Boykov, O. Veksler, and R. Zabih, “Fast Approximate Energy Minimization via Graph Cuts,” IEEE Trans.

Pattern Analysis and Machine Intelligence[J], vol. 23, no. 11, pp. 1222-1239, Nov. 2001.

- [6] J. Yedidia, W. Freeman, and Y. Weiss, "Generalized Belief Propagation," Advances in Neural Information Processing Systems[J], pp. 689-695, 2000.

作者简介

徐戈(1979-), 男, 博士, 主要从事 SAR 图像解译与目标识别研究;

黄志强(1979-), 男, 硕士, 研究方向为雷达信号处理, 雷达系统仿真。

一种适合子带LMS的非均匀滤波器组

张 炜 许小东 戴旭初

(中国科学技术大学电子工程与信息科学系, 安徽合肥, 230027)

摘 要: 为提高子带 LMS 型自适应滤波算法的收敛速度, 提出了一种基于功率谱幅度分割的非均匀滤波器组设计方法。该方法依据最大特征值扩散度最小化准则, 对全频带信号的功率谱幅度进行均匀分割, 并在此基础上, 给出了非均匀滤波器组设计方法。计算机仿真实验和分析表明, 同传统设计方法相比, 在子带数目相同的条件下, 该方法能够有效降低子带信号的最大特征值扩散度, 而信号重建性能基本相同。

关键词: 自适应滤波; 非均匀滤波器组; 功率谱; 特征值扩散度; LMS

Non-uniform Filter Banks with Subband LMS Adaptation

Zhang Wei Xu Xiao dong Dai Xu chu

(Department of Electronic Engineering and Information Science, University of Science and Technology of China, Hefei, Anhui, 230027, China)

Abstract: Based on the magnitude partition with power spectrum of a wideband signal, a non-uniform filter bank was proposed for the purpose of improving convergence speed of the subband LMS adaptation. The proposed method performed an equispaced magnitude partition of the signal PSD following novel criteria of minimizing maximum eigenvalue spread. Accordingly, the structure designing of this class of non-uniform filter banks was provided. Computer simulations and performance analysis demonstrate that, comparing to traditional algorithms with equal number of subbands, the proposed method can efficiently decrease the maximum eigenvalue spread while giving nearly the same reconstruction performance.

Keywords: Adaptive Filtering; Non-uniform Filter Bank; Power Spectrum; Eigenvalue Spread; LMS

1 引言

基于 LMS 型子带自适应滤波是一类重要的自适应滤波方法, 已被广泛应用于宽带信号的自适应处理, 如噪声抵消^[1]、系统辨识^[2]和信道均衡^[3]等场合。子带自适应滤波一般采用滤波器组对信号进行频域子带分解。实现子带分解的滤波器组可分为均匀滤波器组和非均匀滤波器组两大类。已有研究表明, 均匀滤波器组设计较为简单, 但是并没有充分利用信号的特征, 其在降低各子带信号的特征值扩散度以及计算复杂度等方面的性能一般。相比而言, 非均匀滤波器组可根据信号的频谱特点动态地调整滤波器的带宽和中心频率, 因此使用较为灵活, 性能也要优于均匀滤波器组^[5-7]。

非均匀滤波器组设计大都采用了基于合并相邻均匀滤波器的间接设计思想^[8],即首先采用均匀分析滤波器组对全频带信号进行均匀子带分解;然后根据处理带宽内各信号的实际带宽和中心频率,合并若干个相邻的均匀子带信号,完成对全频带信号的非均匀子带分解。应该注意到,这类方法实际只利用了信号的带宽和中心频率信息,通常难以有效降低非均匀子带信号自相关矩阵的特征扩散度,从而导致子带 LMS 型自适应滤波算法的收敛速度较慢。

本文根据 LMS 算法的收敛特性,以提高子带 LMS 型自适应滤波算法的收敛速度为目标,提出了一种设计非均匀滤波器组的新准则和具体实现方法,包括其中的非均匀子带定位、非均匀分析滤波器组和综合滤波器组等关键环节。

2 子带LMS自适应滤波原理

子带 LMS 型自适应滤波算法采用分析滤波器组对全频带的输入信号 $x(n)$ 和误差信号 $e(n)$ 进行分解,得到子带信号 $z_j(n)$ 和子带误差信号 $e_j(n)$, $j = 0, 1, \dots, N-1$, 并在子带中进行滤波器系数的自适应迭代更新,其结构如图 1 所示。

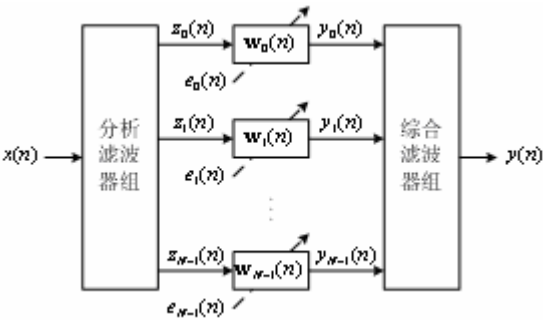


图 1 子带 LMS 自适应滤波结构

图中, $\mathbf{w}_j(n)$ 表示第 j 个子带滤波器的抽头系数矢量, 其自适应更新过程可表示为

$$\mathbf{w}_j(n+1) = \mathbf{w}_j(n) + \mu_j \mathbf{z}_j^*(n) e_j(n), \quad j = 0, 1, \dots, N-1 \quad (1)$$

式中, $\mathbf{z}_j(n)$ 是子带信号 $z_j(n)$ 的矢量形式; μ_j 是步长参数。

已有研究表明, 子带 LMS 型自适应滤波算法的收敛速度快于传统的全频带 LMS 型自适应滤波算法。事实上, 其迭代收敛速度与子带信号 $z_j(n)$ 的自相关矩阵的特征值扩散度 $\bar{X}(\mathbf{R}_{z_j})$ 密切相关, 且全频带收敛速度由收敛速度最慢的子带决定^[9-10]。鉴于此, 如果设计的分析滤波器组能够使最大的子带特征值扩散度最小化, 则可以有效保证等效的子带 LMS 滤波器具有更快的收敛速度。这是本文设计非均匀滤波器组所遵循的基本准则。

3 非均匀滤波器组设计

3.1 功率谱幅度分割思想

令 $\phi_x(\omega)$ 是 $x(n)$ 的功率谱的对数形式, 则由文献[4]可知, $\bar{X}(\mathbf{R}_x)$ 的对数形式可近似为

$$X(\mathbf{R}_x) = 10 \log_{10} \bar{X}(\mathbf{R}_x) \approx \phi_{x, \max} - \phi_{x, \min} \quad (2)$$

其中 $\phi_{x,\max} = \max_{\omega} \{\phi_x(\omega)\}$, $\phi_{x,\min} = \min_{\omega} \{\phi_x(\omega)\}$ 。该式提示我们, 如果 $x(n)$ 的功率谱在通带内有较大的起伏变化, 则 $X(\mathbf{R}_x)$ 较大, 此时 LMS 型自适应滤波算法收敛速度较慢。因此, 可依据全频带信号的功率谱幅度特性, 通过对其进行均匀分割, 实现非均匀子带分解。

图 2 展示了该思想的实现过程。实线表示的曲线是全频带信号 $x(n)$ 的功率谱 $\phi_x(\omega)$ 。不失一般性, 假设将 $\phi_x(\omega)$ 的幅度均匀分割成 4 等份, 即分割成 $\mathbf{r}_j: \phi_j \sim \phi_{j+1}$, $j = 0, 1, 2, 3$, 每个幅度段的长度都为 $\Delta\phi = \phi_{j+1} - \phi_j = (\phi_4 - \phi_0)/4$ 。我们可通过控制功率谱幅度被分割的段数, 来直接控制子带信号的特征值扩散度, 从而为改善和提高子带 LMS 型自适应滤波算法的收敛速度提供了一个有效途径。

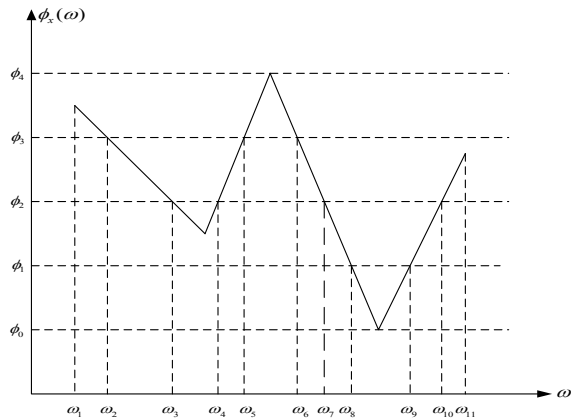


图 2 功率谱幅度分割的原理示意图

幅度的均匀分割导致信号的频带被分割成 10 个子带, 即 $\Omega_k: \omega_k \sim \omega_{k+1}$, $k = 0, 2, \dots, 9$, 每个子带的带宽 $\Delta\omega_k = \omega_{k+1} - \omega_k$ 。注意到, 它们可能并不相等。根据式(2), 与 \mathbf{r}_j 相对应的每个子带都具有相同的特征值扩散度。将特征值扩散度相同的、频域上间隔分布的多个子带信号合并为非均匀子带, 则合并后的非均匀子带数目与功率谱幅度被分割的段数相同。

这种分割方式, 可以灵活地控制子带信号的特征值扩散度, 为实现非均匀滤波器组提供了一种新的有效设计准则, 即最大子带特征值扩散度最小化准则。

3.2 非均匀滤波器组结构

图 3 是基于功率谱幅度分割思想的 LMS 型非均匀子带自适应滤波算法的实现结构。该结构主要包括: 功率谱估计、非均匀子带定位、非均匀分析滤波器组、非均匀综合滤波器组以及子带滤波器系数更新等模块。由于信号功率谱估计已有了很多成熟算法^[11], 而且 LMS 自适应算法及其分析也有很多专著和文献, 所以本文后续只讨论与非均匀滤波器组有关的环节, 即非均匀子带定位算法、非均匀分析滤波器组和综合滤波器组的设计实现。

3.3 非均匀子带定位

假设功率谱幅度被均匀分割成 N 段, 且 $N < M$, 则非均匀子带定位主要是根据功率谱估计的结果, 确定与每个分段 \mathbf{r}_j ($j = 0, 1, \dots, N-1$) 相对应的均匀子带数目和位置 (或编号)。

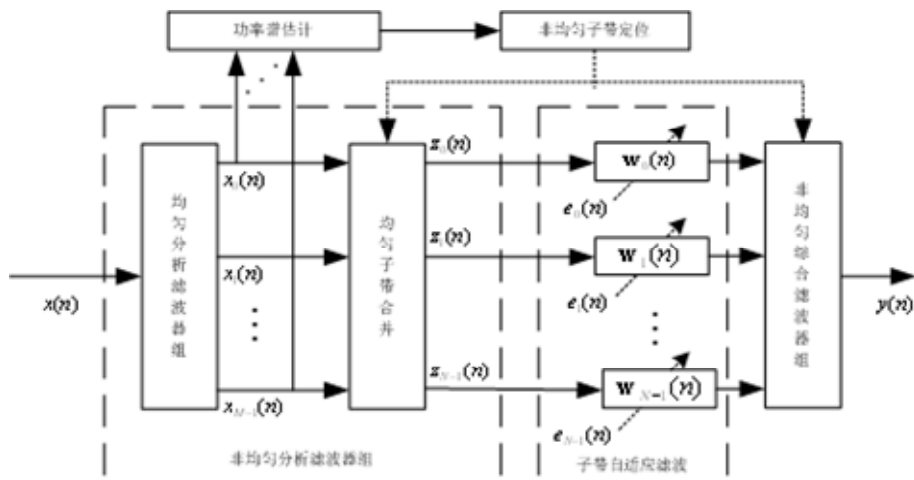


图3 基于功率谱幅度分割的 LMS 型非均匀子带自适应滤波器的结构

在该模块中，首先计算各均匀子带信号和全频带信号的功率谱幅度的取值范围 r_x^i 和 Φ ，并对 Φ 进行 N 等份均匀分割，得到各段的取值区间 r_j 。然后，将 r_x^i 与 r_j 进行比较，如果 r_x^i 包含在某个 r_j 中，则将第 i 均匀子带直接定位到非均匀子带集合 G_j 中。如果 r_x^i 需要若干个 r_j 才能覆盖，则按照非均匀子带的最大特征值扩散度最小化准则对均匀子带 i 进行定位；根据前面的分析，特征值扩散度最大的子带收敛速度最慢，而等效的全频带滤波器的收敛速度由该子带的收敛速度决定，所以采用该准则能够保证全频带等效滤波器的收敛速度最快。

3.4 非均匀分析滤波器组

接下来，我们将利用非均匀子带定位算法提供的均匀子带的位置信息，采用两步合并的方法，设计一种新的非均匀滤波器组，包括分析滤波器组和综合滤波器组。

非均匀分析滤波器组的结构如图 4 所示，由均匀子带分解、频移合并和降采样率三部分组成。令 $H(e^{j\omega})$ 是低通原型滤波器 $h(n)$ 的频率响应，则采用均匀 DFT 滤波器组即可实现对全频带信号 $x(n)$ 进行子带分解，得到均匀子带信号 $x_i(n)$ 。图 4 中，均匀分析滤波器组以多相结构实现， $E_i(z^M)$ 是 $h(n)$ 的 I 型多相成分。根据非均匀子带定位算法的结果 G_j ，完成对均匀子

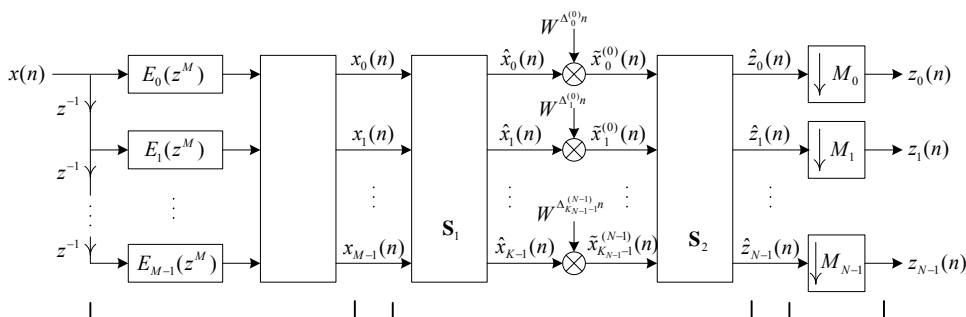


图4 非均匀分析滤波器组

带信号 $i, i \in G_j$ 的频移和直接合并, 即可得到非均匀子带信号 $z_j(n)$ 。值得指出的是, 我们可以构造合并矩阵 \mathbf{S}_1 和 \mathbf{S}_2 来实现合并过程。通过调整矩阵 \mathbf{S}_1 , \mathbf{S}_2 和非相邻子带间距 $\Delta_k^{(j)}$ 的值, 可以简单灵活地合并频域上间隔分布的均匀子带。最后对非均匀子带信号进行降采样处理, 也可以进一步减少算法的计算复杂度, 提高运算效率。

3.5 非均匀综合滤波器组

图 5 是非均匀综合滤波器组结构, 包括升采样、频移复位和均匀子带合成三部分。图中, \mathbf{S}_1^T 和 \mathbf{S}_2^T 是矩阵 \mathbf{S}_1 和 \mathbf{S}_2 转置, M_j , \mathbf{S}_1 , \mathbf{S}_2 和 $\Delta_i^{(j)}$ 的值与非均匀分析滤波器组相同。非均匀综合滤波器组通过升采样和频移复位, 将组成 $z_j(n)$ 的各个均匀子带信号恢复到其实际的频率位置上, 再用均匀综合滤波器组进行全频带信号的重建。均匀子带合成是以多相结构实现的均匀综合滤波器组, $G_i(z^M)$ 是原型滤波器 $f(n)$ 的 II 型多相成分, 各均匀综合滤波器的频率响应为

$$F_i(e^{j\omega}) = F(e^{j\omega} W_M^i) = H^*(e^{j\omega} W_M^i) = H_i^*(e^{j\omega}), \quad i = 0, 1, \dots, M-1 \quad (3)$$

于是, 非均匀综合滤波器组的输出 $y(n)$ 的频域表示近似为

$$Y(e^{j\omega}) \approx \sum_{k=0}^{K-1} \hat{X}_k(\omega) \left[\sum_{i=l_k}^{l_k+n_k-1} F_i(e^{j\omega}) \right] \quad (4)$$

其中, K 为相邻子带组数, l_k 表示第 k 组相邻子带的最小编号, n_k 表示该组相邻子带的个数, 它们的具体数值由定位算法的结果给定, 并且满足 $\sum_k n_k = M$ 。

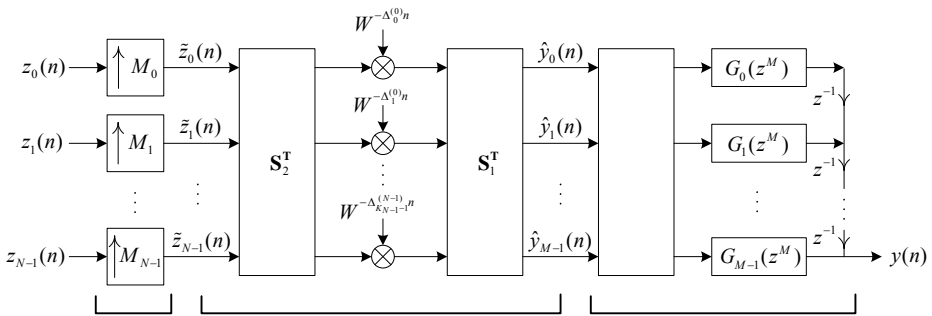


图 5 非均匀综合滤波器组

4 仿真实验与分析

本节利用仿真实验验证和分析所提出的非均匀子带定位算法、非均匀滤波器组设计的有效性以及性能。仿真实验分成两部分: 首先验证非均匀子带定位算法, 并与子带数目相同的均匀子带分解在特征值扩散度方面的性能进行比较; 然后, 检验非均匀滤波器组的重建特性。

仿真实验条件为: (1) 功率谱 $\phi_x(\omega)$ 的幅度被均匀分割的段数 $N = 4$; (2) 在设计非均匀分析滤波器组中, 均匀子带分解的数目 $M = 16$; (3) 原型滤波器 $h(n)$ 是具有线性相位特性的升余弦低通滤波器, 滚降因子 $\alpha = 0.1$, 阻带衰减 δ_1 约为 -45 dB, 通带波动 δ_2 为 -23 dB, 它的幅频响应如图 6 中实线所示; (4) 为了性能比较, 采用均匀分析滤波器组对全频带信号 $x(n)$ 进行均匀子带分解时, 其子带数目也取为 4, 原型滤波器 $h'(n)$ 的带宽是 $h(n)$ 带宽的 4 倍, 其

它的特性参数与 $h(n)$ 相同, 幅频响应如图 6 中虚线所示。(5) 全频带信号 $x(n)$ 的功率谱如图 7 所示。

4.1 非均匀子带定位

采用 Welch 方法^[11]计算各均匀子带信号 $x_i(n)$ 的功率谱估计 $\phi_x^i(\omega)$, $i=0,1,\cdots,15$, 图 7 中相邻两个垂直虚线之间的曲线即为 $\phi_x^i(\omega)$; 按照归一化频率从低到高的顺序, 给 16 个均匀子带从左到右依次编号为 $0,1,\cdots,15$ 。

根据非均匀子带定位算法, 得到 4 个非均匀子带所包含的均匀子带的编号集合分别为: $G_0=\{0,5,6,7,14,15\}$, $G_1=\{1,4,8,13\}$, $G_2=\{2,3,9,11,12\}$, 和 $G_3=\{10\}$; N_j 依次为 6, 4, 5, 1。

依据仿真实验, 全频带信号 $x(n)$ 的特征值扩散度为 30.2dB。采用原型滤波器为 $h'(n)$ 的均匀滤波器组对全频带信号 $x(n)$ 在频域上进行均匀分割后, 4 个子带信号自相关矩阵的特征值扩散度分别为 18.46dB, 19.45dB, 24.06dB 和 29.42dB。而采用对全频带信号 $x(n)$ 的功率谱 $\phi_x(\omega)$ 的幅度进行分割的方法, 并利用非均匀子带定位算法, 所得到的 4 个非均匀子带信号自相关矩阵的特征值扩散度为 7.61dB, 8.12dB, 8.20dB, 10.40dB。通过比较可知, 本文提出的非均匀子带的分解思想和定位算法, 使得 4 个子带信号自相关矩阵的特征值扩散度明显降低, 特别是最大的特征值扩散度由 29.42dB 降低至 10.40dB。

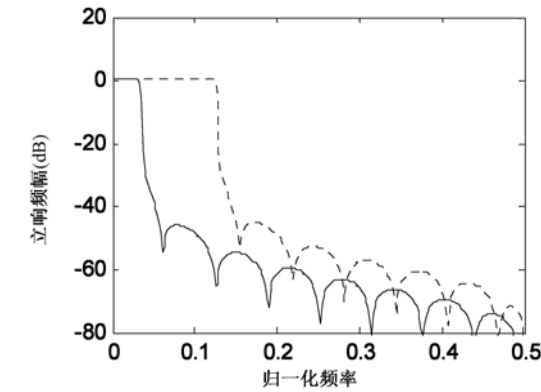


图 6 原型滤波器的幅频响应

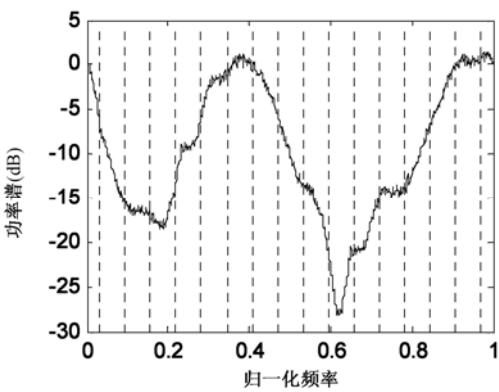


图 7 全频带信号的功率谱

4.2 非均匀滤波器组的重建特性

失真函数是衡量滤波器组重建特性的重要指标。依式(4), 可近似得到失真函数如下

$$T(e^{j\omega})=1-\sum_{k=0}^{K-1}\left[H(e^{j\omega}W_M^{l_k-1})H^*(e^{j\omega}W_M^{l_k})+H^*(e^{j\omega}W_M^{l_k-1})H(e^{j\omega}W_M^{l_k})\right] \tag{5}$$

其中, 该式右端的求和项为失真项, 其主要是由合并非相邻子带导致的。很明显, 只有当 $H(e^{j\omega})$ 满足理想低通特性时, 失真项才等于零。

图 8 是非均匀滤波器组的失真函数 $|T(e^{j\omega})|$ 的曲线, 其中实线表示式(5)的理论分析结果, 虚线表示仿真试验的结果。从图 8(a)中可看出, 理论分析和仿真实验结果非常吻合; 由合并非相邻子带导致的最大幅度失真约为 0.8dB, 这表明本文设计的非均匀滤波器组具有较好的重建特性。图 8(b)是图 8(a)的局部放大图, 可以看到, 理论分析结果在通带内的波动为 0.004dB, 仿真实验结果的波动为 0.008dB, 前者的波动之所以较小, 可能的原因是由于在式(5)的求解过

程中进行了近似，忽略了一些失真所致。

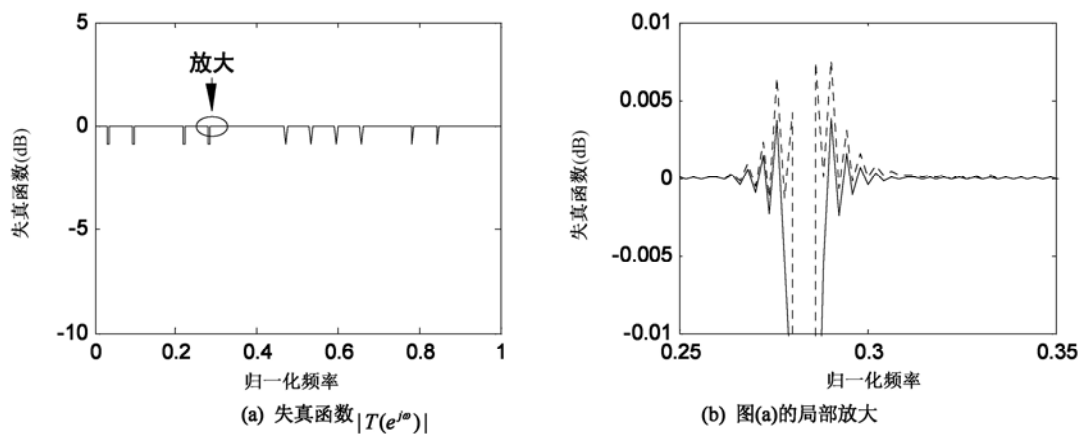


图 8

5 结束语

本文的主要贡献有两点，其一是基于功率谱幅度分割思想提出了设计非均匀滤波器组的一个新准则，即最大的子带特征值扩散度最小化，其二是根据该准则，设计实现了一种非均匀滤波器组，包括非均匀子带定位算法、非均匀分析滤波器组和综合滤波器组设计。计算机仿真实验和分析均表明，同现有的均匀和非均匀滤波器组的设计方法相比，在子带数目相同的条件下，基于本文方法设计的非均匀滤波器组在降低子带信号的最大特征值扩散度方面具有明显优势，而信号重建性能基本相同。

参 考 文 献

[1] Gilloire A, Vetterli M. Adaptive filtering in subbands with critical sampling: analysis, experiments and applications to acoustic echo cancellation[J]. IEEE Trans. on Signal Processing, 1992, 40: 1862-1875.

[2] Schönle M, Fliege N J, Zölzer U. Parametric approximation of room impulse responses by multirate systems[C]. Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing, 1993, I: 153-156.

[3] Weiss S, Dooley S R, Stewart R W, et al. Adaptive equalization in oversampled subbands[J]. Electron. Lett., 1998, 34(15): 1452-1453.

[4] Shynk J J. Frequency-domain and multirate adaptive filtering[J]. IEEE Signal Processing Magazine, 1992, 9: 14-37.

[5] Griesbach J D, Bose T, Etter D M. Non-uniform filterbank bandwidth allocation for system modeling subband adaptive filters[C]. Proc. IEEE Int. Conf. Acoust., Speech, Signal Process., 1999, 3: 1473-1476.

[6] Oakman A, Naylor P. Dynamic structure for non-uniform subband adaptive filtering[C]. IEEE Int. Conf. Acoust., Speech, Signal Process., 2001, 6: 3717-3720.

[7] Griesbach J D, Lightner M R, Etter D M. Constituent subband allocation for system modeling nonuniform subband adaptive filters[J]. IEEE Trans. on Signal Processing, 2005, 53(2): 539-549.

- [8] Cox R V. The design of uniformly and nonuniformly spaced speech coding[J]. IEEE Trans. Acoust., Speech, Signal Processing, 1986, 34:1090-1096.
- [9] Proakis J G 著, 张力军, 张宗橙, 郑宝玉等译. 数字通信 (第四版) [M]. 电子工业出版社, 北京, 2003:478-482.
- [10] Mohamad H, Weiss S, Arif N A M, Alias M Y. Subband decomposition techniques for adaptive channel equalization[C]. 13th IEEE Int. Conf. on Communication, Malaysia, 2005, 2: 617-621.
- [11] Stoica P, Moses R. Spectral Analysis of Signals[M]. Prentice Hall, 2005:41-42.
- [12] Lee J J, Lee B G A design of nonuniform cosine modulated filter banks[J]. IEEE Trans. on Circuits and Systems II: Analog and Digital Signal Processing, 1995, 42(11): 732-737.

3780 点FFT的FPGA设计与实现

陈联武 苏凯雄

(福州大学 物理与信息工程学院, 福建 福州 350002)

摘要: 本文介绍 3780 点 FFT 处理器的设计与实现方法。通过综合利用混合基、素因子和 WFTA 三种算法来完成 3780 点 FFT 的算法设计, 同时在有限字长效应分析的基础上, 提出了一种硬件实现中的移位规则。最后对系统进行了仿真验证, 结果表明其输出信噪比符合 TDS-OFDM 系统的要求。

关键词: 信号与信息处理; FFT; 系统科学法; FPGA

The Design and Realization of 3780-Point FFT Using FPGA

Chen Lian wu Su Kai xiong

(College of Physics and Information Engineering, Fuzhou University, Fuzhou, Fujian 350002, China)

Abstract: The paper introduces the design and realization of 3780-point FFT based on FPGA. The 3780-point FFT was implementet by using of the Common Factor Algorithms (CFA), Prime Factor Algorithms (PFA) and Winograd Fast Fourier Transform Algorithms (WFTA), the effect of finite word length of 3780-point FFT was analyzed, and a rule of shifting during the hardware realization process was proposed. The simulation and analysis of the whole system prove that it can meet the signal-to-quantization noise ratio requirement of the TDS-OFDM system.

Keywords: Signal and Information Processing; FFT; Systematic Scientific Method; FPGA

1 引言

2006 年 8 月国家标准化管理委员会颁布了中国数字电视地面广播传输系统标准 GB20600-2006, 该标准被业界称为 DMB-TH 标准^[1]。DMB-TH 采用了时域同步正交频分复用 (TDS-OFDM)^[2]调制技术, 是 DMB-TH 标准的核心技术。为了实现 TDS-OFDM 调制技术, 3780 点 FFT 处理模块是必不可少的一部分。本文首先介绍 3780 点 FFT 的算法, 接着在有限字长效应分析的基础上, 提出了一种硬件实现中的移位规则, 最后完成 3780 点 FFT 的 FPGA 设计与实现。

基金项目: 福建省科技厅科技计划重点资助项目 (2009H0028)

2 3780 点FFT算法

本文采用综合分解法^[3]来实现 3780 点 FFT。综合分解法是在综合几种离散快速傅里叶算法的基础上提出的。它结合了混合基算法、素因子算法和 WFTA 算法^[4]的优点，在混合基算法的基础上，利用素因子算法消除级与级之间的旋转因子，同时采用 WFTA 算法减少小点数 DFT 的运算量。此外，为了保证原址运算，本文采用了素因子算法的改进算法^[5]。

根据混合基算法、素因子算法及其改进算法的映射公式可推导出各级运算中 RAM 的读写地址公式，具体如表 1 所示。该方法对应的实现框图如图 1 所示。

表 1 各级读写地址公式

各级 WFTA 运算	读写地址公式
7 点 WFTA	$60 \times (9n1 + 7n2) \bmod 63 + n3$
9 点 WFTA	$60 \times (7n1 + 9n2) \bmod 63 + n3$
5 点 WFTA	$(12n1 + 5n2) \bmod 60 + 60 \times n3$
3 点 WFTA	$(20n1 + 3n2) \bmod 60 + 60 \times n3$
4 点 WFTA	$(15n1 + 4n2) \bmod 60 + 60 \times n3$
输出	$(60n1 + n2) \bmod 3780$

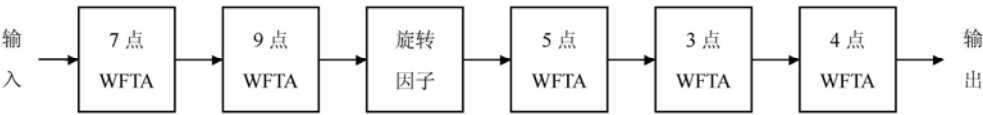


图 1 综合分解法的实现框图

3 有限字长效应分析

在实际硬件设计中，通常采用定点或块浮点运算来实现 FFT。为了设计简便，本文采用定点制来实现 3780 点的 FFT。因此，需要分析算法在硬件实现时产生的有限字长效应，保证系统输出信噪比符合要求。

由 DFT 定义式可知，经过 N 点 FFT，输出数据最大值为输入数据最大值的 N 倍。对于 DMB-TH 系统，其发送端调制输入的数据为经过 QAM 映射的符号，最大值为 18317^[6]。为了保证数据不溢出，运算位宽应为 $\text{lb}18317 \times 3780 = 26.1$ ，取 27 位，这样消耗的硬件资源较多。因此，通过以下分析，可以在保证数据不溢出和运算精度的基础上减少硬件资源消耗。

3.1 输入数据的随机特性

根据 DMB-TH 系统规定，在进入 3780 点 FFT 模块之前，输入码流已经经过了随机化、纠错编码和时频域交织等操作，因此可以认为输入到 3780 点 FFT 模块的符号具有随机特性。利用这一特性，用 Matlab 统计经过 3780 点 FFT 运算后输出数据最大值与输入数据最大值之

比, 可发现输出数据最大值基本上在输入数据最大值的 100~250 倍之间。因此, 为了保证结果不溢出, 运算位宽只需 $1b18317 \times 250 = 22.13$, 取 23 位。

3.2 移位规则

由于 3780 点 FFT 算法的特殊性, 本文接下来提出一种针对该算法硬件实现中的移位规则, 以保证运算过程中的数据不会溢出, 并使得 FFT 输出的信噪比符合系统要求。

在本设计采用的综合分解法中, 将 3780 点 FFT 的计算转化为 7 点、9 点、5 点、3 点和 4 点这五个阶段的 WFTA 运算, 因此, 可以将最后输出与输入数据的倍数分散到每级 WFTA 运算过程中, 再通过对上级 WFTA 运算结果的右移操作来保证下级运算中数据不会溢出, 这样就可以用比较少的运算位宽来实现 FFT 运算。在保证运算过程没有数据溢出的前提下, 通过 MATLAB 统计每级右移的最少位数, 以尽可能保证运算精度。下面以 18 位的运算位宽为例, 说明 3780 点 FFT 硬件实现中的移位规则, 见表 2。

表 2 3780 点 FFT 硬件实现中的移位规则

WFTA	数据位宽	数值范围 (用位数表示)		运算后右移的位数
		运算前	运算后	
7 点	18	16	18	2
9 点	18	$18-2=16$	18	1
5 点	18	$18-1=17$	18	1
3 点	18	$18-1=17$	18	1
4 点	18	$18-1=17$	18	2
输出	16			

根据上述的规则, 用 Matlab 仿真硬件实现中的定点运算过程, 从而确定在硬件实现时的运算位宽、乘法系数和旋转因子的量化位宽。运算位宽 ($N_calculate$)、乘法系数(N_mul)和旋转因子($N_twiddle$)量化位宽与输出信噪比之间的关系如图 2 所示。

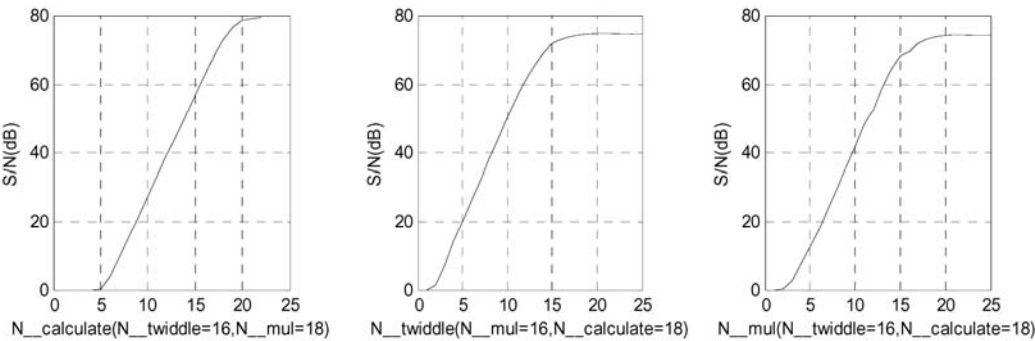


图 2 运算位宽、乘法系数和旋转因子量化位宽与输出信噪比关系

综合考虑硬件资源消耗和系统输出信噪比, 最后取 $N_calculate=18, N_mul=18, N_twiddle=16$, 此时利用 Matlab 计算可得系统输出信噪比为 73.1919dB, 满足 TDS-OFDM 系统的技术指标^[2]要求。

4 硬件实现

4.1 整体设计方案

3780 点 FFT 的硬件实现框图如图 3 所示。该框图包括倍频器，RAM 存储器，选择器，WFTA 运算单元，旋转因子运算单元以及控制单元等。

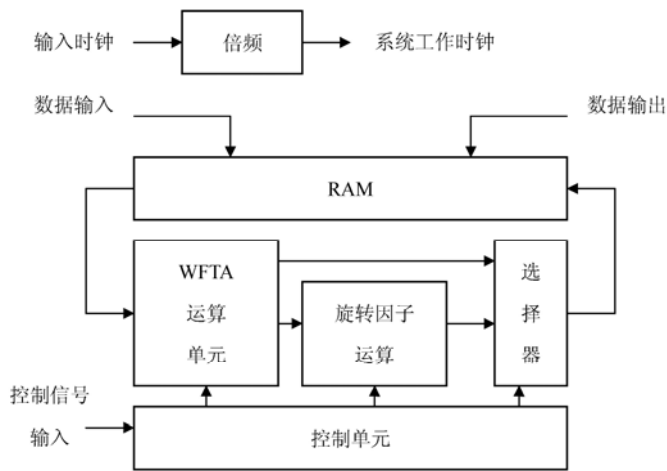


图 3 3780 点 FFT 的硬件实现框图

倍频器是将输入的时钟进行倍频，倍频后的时钟提供给其他模块处理数据使用。控制单元控制各个模块协同工作，每个数据帧到来时启动一次存储操作，每帧的输入数据存储在 RAM 中，等待下个时刻的运算。当运算开始时，把数据从 RAM 中读出，进行 n 点 WFTA 运算，计算完再存到 RAM 的相应地址中，实现原址运算。当进行 9 点 WFTA 运算后，数据需要先乘以旋转因子，然后再通过选择器存入 RAM，而其它点 WFTA 运算后数据直接通过选择器存入 RAM。旋转因子按照一定的顺序预先存在一个 ROM 中。

考虑到输入的符号流一般是连续的，为了保证数据处理的实时性，可以采用基于 3 个存储器的流水线架构，每个存储器大小均为 3780 个符号。从宏观来看，每一个存储器都是在输入、运算和输出三个状态之间进行切换，数据从输入到处理完输出需要两帧的延迟时间。

4.2 仿真实验

系统仿真的测试数据为实部和虚部均为 16 位的随机符号，通过 QuartusII 的时序仿真，结果如图 4 所示。

设计时选用了 CycloneII 系列 EP2C70F672C6 进行时序仿真验证。系统最终消耗 9646 个逻辑单元，573656 个存储单元和 24 个内嵌乘法器。最后系统可支持的最高工作时钟频率为 98.82 MHz，满足 DMB-TH 标准的要求。将仿真输出的数据导入到 Matlab 进行分析，可知系统输出信噪比达到 77dB，符合 TDS-OFDM 系统对信噪比的要求。



图4 3780点FFT的时序仿真结果

5 结论

本文在算法分析的基础上,介绍了3780点FFT处理器的设计和实现方法。在设计过程中对FFT运算的有限字长效应进行分析,提出一种针对3780点FFT硬件实现中的移位规则,在保证系统信噪比的前提下减少硬件资源消耗。系统仿真结果表明,本设计符合DMB-TH标准要求。此外,利用本系统的结构可以很容易地实现由3、4、5、7、9这几个因子组成的FFT,同时为非2的幂次方且能进行因式分解的N点FFT的FPGA设计提供了参考。

参考文献

- [1] 国家标准化管理委员会.GB2060-2006.数字电视地面广播传输系统帧结构、信道编码和调制[S].北京:中国标准出版社.2006-8-18.
- [2] Yang Zhi-xing, HuYu-peng, Pan Chang-yong, et al Design of a 3780-Point IFFT Processor for TDS-OFDM [J]. IEEE Trans, Broadcasting,2002, 48(1): 57-61.
- [3] 杨旭霞,归琳,余松煜.3780点FFT处理器的研究[J].电路与应用,2005,11:32-34.
- [4] 胡广书.数字信号处理—理论,算法与实现[M].北京:清华大学出版社,1997.
- [5] C Sidney Burrus, PETERW Eschenbacher. An In-Place, In-Order Prime Factor FFT Algorithm[J]. IEEE Trans, Acoust, Speech, Signal Processing, 1981, 29(1): 4.
- [6] 陈涛.数字电视地面广播传输系统发端FPGA设计与实现[D].成都:电子科技大学电子工程学院,2007.

作者简介:

陈联武(1986-),男,汉族,福建东山,学士,福州大学物理与信息工程学院,主要从事于数字电视前端系统研究;

苏凯雄(1959-),男,汉族,福建罗源,硕士,教授,博士生导师,福州大学物理与信息工程学院,主要从事无线通信、卫星通信和数字电视系统的研究。

一种基于D-S证据理论的高光谱图像分类方法

佘二永

(中国国防科技信息中心, 北京 100048)

摘要: 提出了一种基于证据理论的高光谱图像分类方法, 通过将高光谱数据划分为多源数据, 然后独立进行各自的分类决策, 最后将来自不同数据源的决策利用证据理论进行融合。同时, 针对典型证据理论不能有效组合冲突证据的问题, 采用了一种基于预处理模式的决策融合方法。最后的分类实验表明提供了比典型证据理论更好的分类结果。

关键词: 高光谱图像; D-S 证据理论; 图像分类; 决策融合

A Classification Algorithm for Hyperspectral Image Based on D-S Evidential Theory

She Er yong

(China Defense Science and Technology Information Center, Beijing 100048, China)

Abstract: A new classification of Hyperspectral data algorithm based on D-S Evidential Theory is proposed in this article. Firstly, the Hyperspectral data is divided into multisource data. Secondly, the local decision is performed at each source data. At last, the set of decisions of each source data is fused in a decision fusion center by D-S evidential theory. But typical D-S theory is sensitive to highly conflict evidences. In this article, a new method based on the pretreatment mode is used. The algorithm presented is applied to real Hyperspectral image data. The experimental result provides better classification accuracy than the method based on typical D-S theory.

Keywords: Hyperspectral Image; D-S Evidential Theory; Images Classification; Decision fusion

1 引言

高光谱遥感数据将传统的图像与光谱信息融为一体, 在获取地表空间图像的同时, 得到每个地物的连续光谱信息^[1]。信息融合技术为高光谱数据的分类提供了新的途径^[2-4], 基于决策融合的方法通过对整个光谱空间进行分解, 将高光谱数据划分为多源数据, 并独立进行各自的分类决策, 然后将来自不同数据源的决策进行融合, 这种处理方法可以有效减小数据维, 避免 Hughes 现象。

证据理论^[5,6]凭借其能够很好地表示“不确定性”、“未知”等概念, 已被广泛地应用于各种数据融合系统中。D-S 证据理论用于多源数据融合时, 由各数据源获得信息产生对目标类别的度量, 由此构成了该理论中的证据, 融合实质是将不同的证据体合并成一个新的证据体的过程。图 1 给出了一种基于 D-S 证据理论的高光谱数据分类流程。

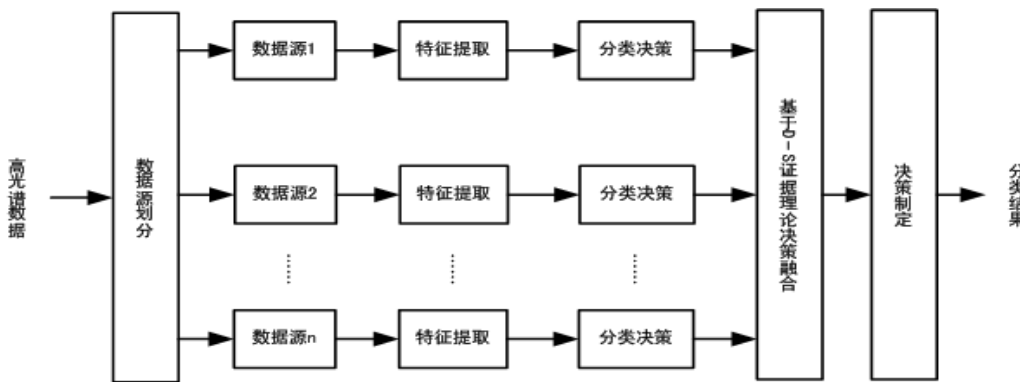


图1 基于 D-S 证据理论的高光谱数据分类流程

2 D-S证据理论的基本原理

2.1 D-S证据理论的基本概念

设 Θ 表示 X 所有可能取值的一个论域集合，且所有在 Θ 内的元素间是互不相容的，则称 Θ 为 X 的识别框架。

定义 1: 假设 Θ 为一识别框架，它的子集的全体即 Θ 的幂集记为 2^Θ ， m 是 Θ 的子集的一个度量，满足：

- 1) $\forall A \subset \Theta, 0 \leq m(A) \leq 1$
- 2) $m(\phi) = 0$
- 3) $\sum_{A \subset \Theta} m(A) = 1$

时，称 $m: 2^\Theta \rightarrow [0, 1]$ 为基本概率分配函数（Basic Probability Assianment, BPA）。 $m(A)$ 表示了对命题 A 的精确信任程度，表示了对 A 的直接支持。

定义 2: 假设 Θ 为一识别框架， $m: 2^\Theta \rightarrow [0, 1]$ 是 Θ 上的基本概率分配函数，定义函数：

- 1) $bel: 2^\Theta \rightarrow [0, 1]$
- 2) $\forall A \in \Theta, bel(A) = \sum_{B \subseteq A} m(B)$

称该函数是 Θ 上的信任函数。显然有 $bel(\phi) = 0$ 和 $bel(\Theta) = 1$ 。信任函数表示命题成立的最小的不确定性支持程度，也称为下限函数。

定义 3: 若识别框架 Θ 的一个子集为 A ，具有 $m(A) > 0$ ，则称 A 为信任函数 bel 的焦元（focal element）。

定义 4: 假设 Θ 为一识别框架，定义 $pls: 2^\Theta \rightarrow [0, 1]$ 为：

$$pls(A) = 1 - bel(\bar{A}) = \sum_{B \cap A \neq \phi} m(B)$$

pls 称为似真函数（plausibility function），表示命题成立的最大的不确定性支持程度，也称为上限函数。

定义 5: $[\text{bel}(A), \text{pls}(A)]$ 称为焦元 A 的证据区间 (evidential interval), 如图 2 所示。

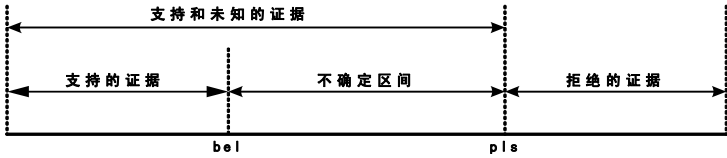


图 2 证据区间示意图

2.2 Dempster 组合规则

证据理论中的 Dempster 组合规则提供了组合两个证据的规则。假设 m_1 和 m_2 是同一辨识框架中的两个相互独立的基本概率分配函数, 用正交和 $m = m_1 \oplus m_2$ 表示它们的组合作用。

定义 6: 设 bel_1 和 bel_2 是同一识别框架 Θ 上的两个信任函数, m_1 和 m_2 分别是其对应的基本概率赋值, 焦元分别为 A 和 B , 则:

$$m(C) = \begin{cases} 0, & \text{if } C = \phi \\ \frac{\sum_{A \cap B = C} m_1(A)m_2(B)}{1 - K} \end{cases} \quad (1)$$

其中, K 表示了证据冲突 (Evidential Conflict) 的额度:

$$K = \sum_{A \cap B = \phi} m_1(A)m_2(B) \quad (2)$$

若 $K \neq 1$, 则 m 也是一个基本概率分配; 若 $K = 1$, 则表明 m_1 和 m_2 互相矛盾, 不能对基本概率赋值进行组合。对于多个证据的组合, 可采用定义 6 给出的组合规则对证据两两组合。

3 数据源划分

对于高光谱数据空间最直接的划分方法就是进行等间隔划分, 这种划分虽然实现起来最简单, 但不能有效去处数据源之间的相关性。由于高光谱数据覆盖可见光到红外的光谱范围, 利用光谱段的差异也可以进行数据源的划分。这种划分方法相对来说具有合理的划分依据, 但是没有充分考虑到高光谱数据对应的场景不同以及地物类型的变化对数据之间的关系所造成的影响^[7]。

通过对高光谱数据的相关性分析, 如图 3 给出的全局相关矩阵, 可以看出高光谱数据波段之间相关性相差很多, 根据全局相关性来划分数据源是有效的。当两个波段间的相关性满足给定的阈值参数时, 这两者之间的波段便构成一个独立的数据源, 每个独立的数据源中的波段都具有类似的相关性。在划分数据源完成后, 由于数据源内的波段相关性较强, 而各个数据源之间的波段相关性较弱, 因此在后续的处理中可以将各个数据源近似看作是彼此相对独立^[3, 7, 8]。

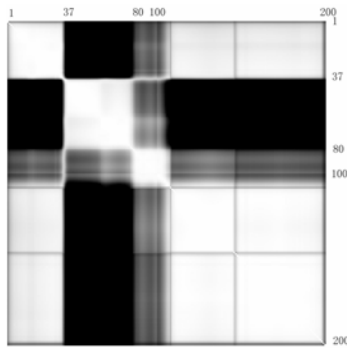


图3 高光谱数据全局统计相关矩阵（黑色为0，白色为1）

4 特征提取

主分量分析（PCA）是一种广泛使用的特征提取技术，其主要思想是将原始 n 个波段中的所有信息压缩到 P 个新波段中，以消除波段间的相关冗余。具体步骤如下：

- 1) 计算数据源的协方差矩阵 Σ 。
- 2) 计算矩阵 Σ 的特征值 λ_r 和对应的特征向量 $\phi_r (r=1, 2, \dots, n)$ 。
- 3) 将特征值 λ_r 按由大到小的次序排列，即

$$\lambda_1 > \lambda_2 > \dots > \lambda_n \quad (3)$$

- 4) 选择前 p 个特征值对应的 p 个特征向量构造变换矩阵 H_p 。
- 5) 进行变换得到的新模式 f' ，变换公式为：

$$f' = H_p f \quad (4)$$

原始数据经过 PCA，其方差分布将主要集中在前 p 个主成分，这些特征含有的信息量最大。采用主分量特征提取方法的一个关键的问题是维数的确定，根据特征值来选择主成分时，由于不同数据情况的特征值分布不同，其特征维数也不同，我们根据特征值的累积百分比 V_r 来选取主成分的数目， V_r 的定义如下：

$$V_r = \frac{\sum_{i=1}^p \lambda_i}{\sum_{i=1}^n \lambda_i} \quad (5)$$

其中 λ_i 为第 i 个特征值， p 为保留的特征值个数， n 为数据源的波段数。

5 基本概率分配函数的构造

由 D-S 证据理论进行融合，首要问题是构造相应数据源的基本概率分配函数。在证据理论中，确定基本概率分配函数没有统一的办法，基本概率分配函数只需要满足式（5）给出的简单条件。但在工程中，基本概率分配函数一般满足如下的约束，即除去 Θ 之外，其余焦元相交为空。基于此约束条件，我们采用一种由似然函数来构造相应基本概率分配函数的简单

方法。

假设识别框架 $\Theta = \{\theta_1, \theta_2, \dots, \theta_m\}$ ，对于一个分类样本 x ，类别 θ_i 的似然函数为 $p(x|\theta_i)$ 。由此引入基本概率赋值函数：

$$m(\theta_i) = \alpha \times \frac{p(x|\theta_i)}{\sum_{j=1}^m p(x|\theta_j)} \quad (6)$$

$$m(\Theta) = 1 - \alpha \quad (7)$$

其中 $\alpha \in [0,1]$ 为数据源的分类性能参数，可以由先验知识、数据源的分类精度、以及采用数据源的可分性判据来给定。

6 基于预处理模式的决策融合方法

由于 Dempster 组合规则的主要问题在于冲突证据之间的组合，而对于非冲突证据之间的组合是成功的。因此，我们采用了一种基于预处理模式的证据理论改进方法[9]，其主要思想是在采用 Dempster 组合规则进行证据组合之前，通过将冲突焦元的基本概率赋值部分转移到焦元并集，有效的将冲突证据变为非冲突证据。

假设 A_m 和 B_n 分别是证据 m_1 和 m_2 的焦元，则证据 m_1 通过预处理分配基本概率给 $A_m \cup B_n$ 的公式如下：

$$m_1^C(A_m) = (1 - w_{(A_m, B_n)})m_1(A_m) \quad (8)$$

$$m_1^C(A_m \cup B_n) = w_{(A_m, B_n)}m_1(A_m) \quad (9)$$

其中 $w_{(A_m, B_n)}$ 为针对证据 m_1 和 m_2 的冲突焦元 A_m 和 B_n 的权值，由于权值表现了冲突程度的不一致性参数。因此，把权值定义为：

$$w_{(A_m, B_n)} = \begin{cases} m_1(A_m) \times m_2(B_n); & \text{当 } A_m \cap B_n = \Phi \\ 0; & \text{当 } A_m \cap B_n \neq \Phi \end{cases} \quad (10)$$

可知经过预处理后的基本概率赋值符合基本概率分配函数的定义（定义1）。

由上式可知，预处理前冲突焦元 A_m 和 B_n 的冲突额度 K ：

$$K = m_1(A_m)m_2(B_n) \quad (11)$$

而经过预处理后的冲突额度 K^C 为：

$$\begin{aligned} K^C &= (1 - m_1(A_m)m_2(B_n))^2 m_1(A_m)m_2(B_n) \\ &= (1 - K)^2 K \end{aligned} \quad (12)$$

可见，预处理后的冲突额度大大降低。

并且，当冲突焦元 A_m 和 B_n 的冲突额度 $K \rightarrow 1$ 时，预处理后的基本概率分配为：

$$m_1^C(A_m \cup B_n) \rightarrow m_1(A_m), \quad m_1^C(A_m) \rightarrow 0 \quad (13)$$

此时，冲突部分的基本概率赋值大部分转移了冲突焦元的并集。而当冲突焦元 A_m 和 B_n 的冲突额度 $K \rightarrow 0$ 时：

$$m_1^C(A_m \cup B_n) \rightarrow 0, \quad m_1^C(A_m) \rightarrow m_1(A_m) \quad (14)$$

此时，冲突部分的基本概率赋值仍主要在焦元本身上进行分配。

由于经过预处理后容易产生过多的基本概率赋值较小的焦元，这将使组合的计算量增加，对于这个问题，可以采用通过设置门限的方法对冲突较大的焦元进行预处理，即：

$$w_{(A_m, B_n)} = \begin{cases} w_{(A_m, B_n)}; & \text{当 } w_{(A_m, B_n)} \geq T \\ 0; & \text{当 } w_{(A_m, B_n)} < T \end{cases} \quad (15)$$

由于预处理方法不满足结合律的要求，这将使组合结果依赖于证据组合顺序。因此，实际应用中先组合相互间冲突额度大的证据，然后再组合冲突额度小的证据。

7 决策制定

决策制定主要是根据由证据理论所得到的信度区间进一步推出最终的决策，使证据区间能真正用于识别决策的制定。可以见到的决策制定方法有：基于最大信任函数的方法、基于最大似真函数的方法、基于信任函数的“最少点”原则的方法[5]、基于基本概率赋值的方法[10]、基于最小风险的方法[10]、基于证据区间的方法[11]、基于 Pignistics 概率的方法[12, 13]。其中基于证据区间的方法利用基数函数对命题的不确定区间重新进行分配将命题的不确定性描述由区间转变为一个数值，这种决策方法比较符合人类的直觉，公式如下：

$$uu(\theta_i) = \sum_{A_j \in \Theta} \frac{\text{Card}(\theta_i \cap A_j) \times m(A_j)}{\text{Card}(A_j)} \quad (16)$$

其中， $uu(\theta_i)$ 定义为对类别 θ_i 的不确定性度量， $\text{Card}(\bullet)$ 是预先定义的基数函数，通常可取值命题所含元素个数。并且当 $\text{Card}(\bullet)$ 取值命题所含元素个数时，公式等同于 Smets 提出的基于 Pignistics 概率的决策方法[12]。

8 实验结果与分析

我们利用典型的 AVIRIS 遥感高光谱图像进行了分类试验[14]。原始图像取自 1992 年 6 月拍摄的美国印第安纳州西北部印第安松树试验区的一部分。去掉二氧化碳和水气吸收带（104—109、150—163），从原始的 220 个波段中选取了 200 个波段作为研究对象，每个波段图像的大小为 145×145 ，图 4 给出了假彩色合成图像。

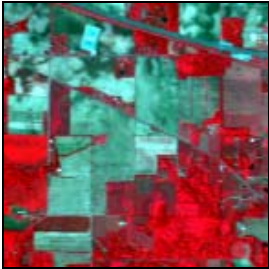


图 4 高光谱假彩色合成图像

根据图 3 给出的全局统计相关矩阵，可以有效的将高光谱图像划分为多源数据，数据源

划分如表 1 所示。

表 1 数据源划分

	波段划分	波段数目
数据源 1	1-36 波段、100-103 波段、110-149 波段、164-220 波段	137
数据源 2	37-79 波段	43
数据源 3	80-99 波段	20

实验中我们选用了 4 类地物：大豆、玉米、草地和林地。各地物的训练、测试样本数目分布如表 2 所示。

表 2 样本分布

	大豆	草地	玉米	林地
训练样本数目	144	200	121	231
测试样本数目	590	480	1183	863

试验选取特征值的累计百分比 $V_r = 0.995$ ，各种数据源特征选取数目如表 3 所示。虽然特征提取不可避免地存在信息的丢失，甚至会造成分类性能的降低。但是应当指出，变换后参与分类的特征数减少，缩短了分类时间，降低了分类的复杂性。

表 3 数据源特征选取数目

	数据源 1	数据源 2	数据源 3
特征选取数目	19	4	7
V_r	0.99529	0.99574	0.99557

试验中根据数据源的预分类来决定数据源权值的大小，先对各数据源进行预分类，根据预分类的精度来决定它们对最终结果的贡献，并以此来分配权值。图 5～图 8 分别给出了采用真实类别标记结果、采用 Bayes 融合分类结果、采用标准 D-S 证据理论融合分类结果和采用预处理的 D-S 证据理论融合分类结果。从分类结果图像中可以看到，通过预处理可以达到较好的分类效果。

表 4 分类结果比较

	训练样本 1	测试样本
Bayes 融合分类	99.7%	67. 68%
D-S 融合分类	99.8%	68. 27%
D-S 预处理融合分类	99.8%	69. 34%

从表 4 中给出的样本数据和测试数据的分类精度比较也可以看出，通过预处理方法进行 D-S 证据融合可以有效提高分类精度。



图 5 真实类别标记



图 6 Bayes 融合分类结果



图 7 典型 D-S 融合分类结果

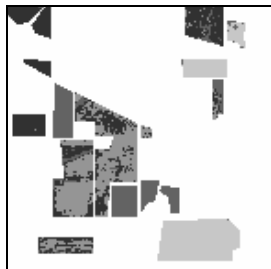


图 8 改进方法融合分类结果

9 小结

本文提出了一种基于证据理论的高光谱图像分类方法，将高光谱数据划分为多源数据，并独立进行各自的分类决策，然后将来自不同数据源的决策利用基于预处理模式的证据理论进行融合。真实的多光谱图像分类实验表明提供了更好的分类结果。

参 考 文 献

- [1] 浦瑞良, 宫鹏, 高光谱遥感及其应用[M]. 2000: 高等教育出版社.
- [2] A.Wilson, Perceptual-based image fusion for hyperspectral data[J]. IEEE Trans. Geosci. Remote Sensing, 1998. 35(4): 1007-1017.
- [3] Jon Atli Benediktsson, Ioannis Kanellopoulos, Classification of Multisource and Hyperspectral Data Based on Decision Fusion[J]. IEEE Trans. Geosci. Remote Sensing, 1999. 37(3): 1367-1377.
- [4] L.O.Jimenez, A.M.Morell, A.Creus, Classification of Hyperdimensional Data Based on Feature and Decision Fusion Approaches Using Projection Pursuit, Majority Voting, and Neural Networks[J]., IEEE Trans on Geoscience and Remote Sensing 1999. 37(3): 1360-1366.
- [5] 段新生, 证据理论与决策、人工智能[M]. 1993: 北京:中国人民大学出版社.
- [6] G.Shafer, A Mathematical Theory of Evidence[M]. 1976: Princeton: Princeton University Press.
- [7] 张钧萍, 张晔, 周廷显, 基于信息融合的超谱图象分类方法研究[J]. 哈尔滨工业大学学报, 2002. 34(4): 464-468.
- [8] I.A.Benediktsson, P.H.Swain, Consensus Theoretic Classification Methods[J]. IEEE Trans on Geoscience and Remote Sensing, 1992. 22(4): 688-704.
- [9] 余二永, 王润生, 徐学文, 基于预处理模式的 D-S 证据理论改进方法[J], 模式识别与人工智能, 2007. 20(5): 711-715

- [10] 何友, 王国宏等, 多传感器信息融合及应用[M]. 2000: 北京: 电子工业出版社.
- [11] 刘大有, 欧阳继红, 唐海鹰等, 一种简化证据理论模型的研究[J]. 计算机研究与发展, 1999. 36(2): 134-138.
- [12] P.Smets. Constructing the pingnistic probability function in a context of uncertainty[C]. Proc. Uncertainty in Artificial Intelligence. 1990. 29-40.
- [13] J.J.Sudano, L.Martin. Pignistic Probability Transforms for Mixes of Low- and High -Probability Events[C]. Proc. International Conf. on Information Fusion. 2001. 23-27.
- [14] http://dynamo.ecn.purdue.edu/~biehl/MultiSpec/download_win.html.

作者简介:

余二永 (1976-), 男 (汉族), 江苏省连云港人, 博士, 高级工程师, 主要研究方向: 目标特性、图像分析、模式识别和信息融合。

通信地址:北京 122 信箱 6 分箱

邮政编码:100048

联系电话:15911021977

E-mail: ey_s@yeah.net

简化的SMV算法及其仿真

王小兵 黄 冰 杜俊宇

(桂林电子科技大学 信息与通信学院, 广西桂林 541004)

摘 要: 介绍了变速率语音算法 SMV, 详细给出算法改进细节和实现方法。SMV 的改进部分包括信号修正、LSF 量化、固定码本搜索, 采用 C 代码实现改进点, 并在 Visual C++ 6.0 中仿真成功。仿真结果表明: 改进后算法的计算复杂度有了显著的降低, 并且获得的语音质量几乎没有下降。

关键词: 变速率语音算法; SMV; 计算复杂度; 语音质量

Reduction and Improvement for 3GPP2 SMV

Wang Xiao bing Huang Bing Du Jun yu

(Institute of Information and Communication, Guilin University of electronic Technology, Guilin 541004, China)

Abstract: A improving algorithm of SMV was described. Variable-rate speech coding algorithm SMV has high computational complexity. So SMV was inefficient for real time implementation. The improvement of SMV were Signal Modification, LSF Quantization and Fixed Codebook Search, and have been implemented in float-point on the platform of Visual C++ 6.0. The improved SMV reduced its computational complexity and its voice quality was identical to SMV.

Keywords: variable-rate speech algorithm; SMV; computational complexity; voice quality

1 引言

SMV 是由 3GPP2 组织于 2001 年提出的 3G 移动通信标准^[1], 提供更高的通信语音质量和更低的占用带宽。SMV 计算量很大, 需要提高硬件设施性能才能应用于移动通信网络^[2]。

2006 年, 德州仪器的四位工程师 C.V.Goudar、Pankaj Rabha、Murali Deshpande 和 Ajit Rao 提出低复杂度的 SMV 版本^[3]: SMV Lite, 但是所提供的资料仅仅说明了改进的方向, 并没有给出具体细节。3GPP2 要求 SMV 的平均码率和语音质量都要符合 3G 标准, 但是不要求精确的比特实现。在不影响通话语音质量的前提下, SMV 算法得到简化, 降低了运算量。

2 SMV Lite 改进原理

标准 SMV 编码部分模块多, 需要处理话音激活、语音分类、信号修正、码本搜索等。这

些功能模块中，信号修正和 FCB 搜索占用的整体运算量较大，减小其运算量对提高算法实用性，相当有益；LSF 量化的改进也涉及到量化码本搜索，优化运算结构以便减小运算时间，从而快速获得最佳量化码矢量。SMV Lite 的改进部分主要集中在以下三个子模块：信号修正，LSF 量化和 FCB 搜索。SMV Lite 的比特码流能与 SMV 兼容(其编码器产生的比特码流可由 SMV 的解码器解码)。

2.1 信号修正

SMV 的信号修正通过连续地修正加权语音信号使之与基音包络相匹配，以达到增强声门参数的稳定性，进而增强基音周期稳定性的目的。首先，SMV 分析加权语音信号，以获得基音周期脉冲；其次，以基音周期脉冲为中心，将一帧语音分成长度可变的几个子帧；最后，对每个可变子帧进行小数基音周期搜索，获得的精确基音周期值用于语音信号的非线性修正，可以使得低速率激励码的基音包络更加规则。

SMV Lite 继承了 SMV 信号修正的基本原理。由于基音周期搜索和信号修正两者的计算复杂度是与可变子帧的数量成正比的。在基音延时较短的女性声音分析中，将会分出大量的可变子帧。SMV Lite 不参考整帧基音周期值，将可变子帧长度的最小值限制为 12 个样点，这将确保修正区域不会分裂任何一个基音周期的峰值，可以显著地减小最坏情况下的 WMOPS 值。

2.2 LSF 量化

标准 SMV 的 LSF 编码是对预测误差进行多阶矢量量化的结果，并且在所有速率条件下，量化表和量化机制是一样的。将输入的每一语音矢量与码本矢量对比，得出加权的平均平方误差值 WMSE。当 WMSE 最小时，便取得量化矢量。减小 WMSE 的计算时间，快速得出量化矢量，就需要优化 WMSE 计算过程。其 WMSE 的 metric 值为：

$$WMSE_{LSF}^{j,i,r} = \sum_{i=1}^{10} w_{LSF}(i) (I_{LSF}^{j,l}(i) - C_{LSF}^{j,r}(i))^2 \quad (1)$$

其中 j 为预测量化器的阶，l 为输入矢量的索引，i 为线谱对序号， $w_{LSF}(i)$ 为第 i 个线谱对的加权值， $C_{LSF}^{j,r}(i)$ 为第 j 阶量化码本的第 r 个矢量。在标准 SMV 中，代码直接计算公式(1)使之最小，导致较高的计算量。

为了尽快得出 WMSE 值，SMV Lite 建议先做预计算处理，以减小计算复杂度。简单的预计算不会影响浮点运算中 LSF 编码矢量的选择。

2.3 固定码本搜索

SMV 的固定码本由多个子码本组成，对其搜索将产生巨大的运算量。下面将使用三种技术来减小 FCB 搜索的计算复杂度。

2.3.1 简化的后向基音增强

标准 SMV 对脉冲码本进行基因增强，其中后向基音增强提供码本的非因果贡献，且无法合并。当一帧语音分为三子帧且基音周期很小时，后向基音增强增加的运算量非常大，而且它会对子帧中三个基音周期延时的语音造成指数级的幅度衰减。在子帧中这种情况经常发生，

这就需要采用一定的方法限制 FCB 中后向基音增强的使用。

在高运算量的特定环境下(一帧语音分为三子帧且基音周期小于一定门限), SMV Lite 会把这种影响限制在两个基音周期内。通过观察, 在最坏的计算复杂度条件下, 后向基音增强的运算量显著地减少了, 并且对语音质量几乎没有影响。

2.3.2 可选的节点搜索

标准 SMV 在处理类型 1 全速率语音子帧的八脉冲固定码本时, 搜索包括两轮: 第 1 轮, 搜索时添加脉冲的方式是一次一对(共四对)。考虑脉冲对所有可能位置(满搜索), 并在标准加权误差最小时, 取得最佳脉冲位置; 第 2 轮, 脉冲对的位置可以再优化(去掉冗余的脉冲位置), 其余 3 对脉冲依次执行; 然后改为一次优化一个脉冲位置。每一子帧都要经历上面两个过程, 每一轮脉冲的位置都要修改, 必然产生非常大的运算量。

在 SMV Lite 中, 首先, 一次一个脉冲地搜索, 同时脉冲位置依次优化。目的仅仅是为了获得一个最佳脉冲可能位置的估计; 然后, 应用 SJS(Selective Joint Search 可选的节点搜索), 某些轨道上的脉冲很可能对语音质量有很大影响, 仅选取和再优化这些脉冲, 而不是盲目全程优化。轨道的选择是基于标准加权误差最小化。第 2 轮中取代盲目满搜索的策略节省了大量运算, 并且对语音质量没有影响。

2.3.3 简化的码本搜索空间

类型 1 语音帧的 FCB 具有本质特性---周期性, 且仅包含一个二脉冲子码本。SMV 的处理方式是: 首先, 对二脉冲子码本进行预搜索, 选择 16 或 19 个最佳位置; 然后, 对二脉冲子码本进行满搜索。

SMV Lite 在预搜索过程中, 利用周期性可以有效地减小类型 1 语音帧的搜索空间和计算复杂度。由于类型 1 语音帧有较高的基音增益, 每一个待选的 FCB 脉冲实际上被余下的语音子帧在基音增强时覆盖了, 搜索那些最佳脉冲位置以外的其他脉冲位置是多余的。限制码本矢量的搜索空间可以显著地减少 FCB 搜索的运算量。

3 SMV Lite改进算法实现

为了实现这些改进点, 必须对应 SMV 算法的说明文档^{[4][5]} ARIB STD-T64-C.0030-0 V2.0 (下文出现的功能函数和计算公式都包含在此文档中), 修改相应的算法计算公式, 然后在 C 代码^[6]中修改相应的执行语句和运算结构。

3.1 SMV Lite信号修正的实现

信号修正改进算法即是要对可变子帧的长度进行限制, 原子帧长度为 $\min\{T_{\text{peak}} - T_0 + L_{\text{puls}}, L_{\text{frm}} - i_0\}$, 则将原计算式变换为:

$$L_{\text{sub}} = \max\{\min\{T_{\text{peak}} - T_0 + L_{\text{puls}}, L_{\text{frm}} - i_0\}, 12\} \quad (2)$$

其中 L_{sub} 为子帧长度, T_{peak} 为子帧脉冲峰值位置, T_0 为第一子帧起始时刻, L_{puls} 为脉冲

宽度, L_{frm} 为整帧长度, i_0 为当前子帧起始时刻。如果 $T_{peak} + 1.25 \square Pit(i_0) > L_{olpit} - L_{lkhd} + \Delta_{SM}$ (其中 $Pit(i_0)$ 为基因包络, L_{olpit} 为开环基因延时, L_{lkhd} 为 lookahead buffer 的长度, Δ_{SM} 为信号修正累计延时), 这时需要估计下一子帧可能的脉冲峰值位置 T'_{peak} , 若 $T'_{peak} > L_{oppit} - L_{lkhd} + \Delta_{SM}$ (L_{oppit} 为最佳基音延时), 即 T'_{peak} 超过了整帧, 此时当前子帧长度 L_{sub} 的计算式应变化为:

$$L_{sub} = \max \{L_{frm} - i_0, 12\} \quad (3)$$

说明文档中函数 `PPP_pitch_preproc` 的作用是修正加权语音以增强长时预测, 通过运算以确定子帧长度, 修改公式(2)和(3)相应的执行代码, 即可将子帧最小长度限制为 12 个样点。

3.2 SMV Lite 中 LSF 量化的实现

LSF 量化中 WMSE 的 metric 值计算公式可做变形, 简化平方因式, 并且展开成多个因式之和(差), SMV Lite 的处理形式如下:

$$\begin{aligned} WMSE_{LSF}^{j,i,r} = & \sum_{i=1}^{10} w_{LSF}(i)(I_{LSF}^{j,l}(i))^2 + \sum_{i=1}^{10} w_{LSF}(i)(C_{LSF}^{j,r}(i))^2 \\ & - 2 \sum_{i=1}^{10} w_{LSF}(i)(I_{LSF}^{j,l}(i) \square C_{LSF}^{j,r}(i)) \end{aligned} \quad (4)$$

$w_{LSF}(i)$ 和 $C_{LSF}^{j,r}(i)$ 可看作常量, $I_{LSF}^{j,l}(i)$ 仅有为变量矢量。可以清楚地看到, 在初始化分步搜索过程之前, 第二项为常量, 第一项和第三项都能有效地进行预计算和存储中间值。公式(4)中则变为仅对 $\sum_{i=1}^{10} (I_{LSF}^{j,l}(i))^2$ 和 $\sum_{i=1}^{10} (I_{LSF}^{j,l}(i))$ 的计算, 也减小了计算复杂度。函数 `LSF_Q_lsfqnt` 运用滑动平均残差量化算法来量化 LSF, 改变原运算结构, 执行公式(4)计算过程, 即可减小得出结果的时间。

3.3 SMV Lite 中固定码本搜索的实现

3.3.1 简化的后向基音增强实现

对子帧进行后向基音增强时, 首先将从脉冲子码本搜索获得的最优脉冲作为主脉冲, 然后在位于主脉冲之前的基因周期整数倍的位置上插入增强脉冲, 脉冲的符号与主脉冲符号保持一致。其增强系数为:

$$P_{mag}(m) = (\hat{g}_a)^m \quad m=1,2,3,4 \quad (5)$$

其中 \hat{g}_a 为修正后的基因增益, m 为后向基音增强的序号。说明文档中函数 `FCS_PreCalc_FBCB` 执行基因增强的过程, 当 $l_{INT}^p \leq L_{SF} / 4$ 时(l_{INT}^p 为子帧的基因周期延时), 将 m 的取值范围限定为[1,2], 即可将后向基音增强产生的非因果响应的影响控制在两个基因周期延时范围内。

3.3.2 可选的节点搜索实现

SMV Lite 调用函数 `FCS_cdbk_search_8500BSP` 处理类型 1 全速率语音子帧的八脉冲固定码本。用加权输入语音和加权重构语音之间的均方差最小化准则来选取最佳脉冲位置。首先,

让单个脉冲独享一个轨道，而不必与另外一个脉冲组合。调用函数 FCS_Fast_Search_FBCB(单脉冲位置搜索)搜索八个脉冲和所有可能位置，就可以知道八个轨道上哪些位置的脉冲给予加权平方误差的贡献最小，选取这些脉冲位置。这些脉冲位置仅仅只是估计而已，其数目远小于满搜索的位置数目。然后，将八脉冲组合成为四个脉冲对，进行联合搜索，调用函数 FCS_Fast_Search_FBCB2(双脉冲联合位置搜索)，得到最终优化的八个脉冲。这是使用非穷尽搜索策略，以便快速获得局部最优码矢量。

3.3.3 简化的码本搜索空间实现

码本搜索空间的简化是针对类型 1 语音(稳定浊音)的二脉冲子码本。SMV Lite 实现简化的关键是要减少可能脉冲位置的数目。首先，执行预搜索以限制位置数目，其依据为下式：

$$N_{pre} = \begin{cases} 16 & type-1 frame, l_{INT}^p < 30 \\ 19 & type-1 frame, l_{INT}^p \geq 30 \end{cases} \tag{6}$$

其中 l_{INT}^p 为子帧基音延时， N_{pre} 为预搜索后获得的可选脉冲位置个数；这 N_{pre} 个位置的脉冲使得下式误差最大化： $\hat{\epsilon}_i = (b_{Tg} p_i)^2 / \|H p_i\|^2$ (其中 b_{Tg} 为加权滤波的目标信号， p_i 为基音增强系数， H 为合成滤波器冲击响应的下三角矩阵)；其次，仅在这 16 或 19 个最佳位置中进行双脉冲联合位置搜索，而不必要在二脉冲子码本的 52 个位置上进行满搜索；最后，进行单脉冲位置搜索，即可获得所需要的两脉冲位置，以及相应的码矢量。

4 SMV Lite 仿真结果

ARIB(日本无线工业及商贸联合会)官方提供了 SMV 浮点源代码，在其基础上进行大量修改，实现了信号修正、LSF 量化和 FCB 搜索算法的改进，修改后的代码符合 SMV Lite 算法标准。以 Visual C++ 6.0 为仿真平台，运行成功，解码后获得的语音质量较好。运算量由 WMOPS 值表示，表 1 给出了三个子模块算法的改进前后的运算量。

表 1 改进前后的 WMOPS 对比

	信号修正	LSF 量化	固定码本搜索
SMV	6.11	3.26	11.32
SMV Lite	3.65	2.51	7.13
优化比例	40.2%	23%	37%

对同一语音在不同模式下进行编码，得到的平均编码速率符合 3G 标准，如表 2(ADR 单位：kbps)。

表 2 SMV Lite 的 6 种模式下的 ADR

模式	0	1	2	3	4	5
ADR	5.40	4.61	3.42	2.87	3.27	2.75

为了测试 SMV Lite 算法的音质效果，将多种语音由 SMV Lite 处理，典型对比声音为无噪语音、噪声、带噪语音，如图 1 所示。SMV Lite 保留了语音，抑制了噪声。

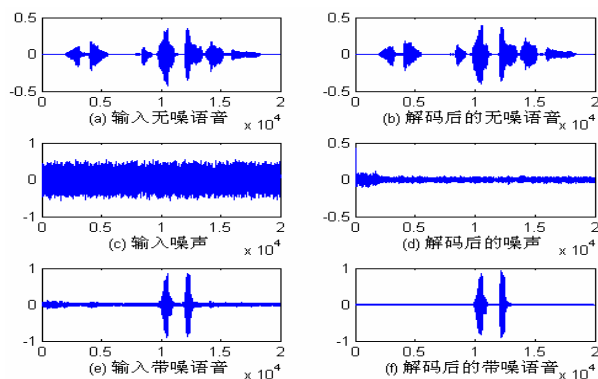


图 1 SMV Lite 处理语音前后的对比

5 结论

SMV Lite 算法的计算复杂度小于 SMV，且获得的语音质量几乎相同。SMV Lite 是应用成本、语音质量和占用带宽的有效折衷。

参 考 文 献

- [1] Yao Gao,Eyal Shlomot,Adil Benyassine et al.“The SMV Algorithm Selected by TIA and 3GPP2 for CDMA Applications”,IEEE[S], Pro ICASSP-2001:PP.709-712
- [2] Erh Wen Hu,Cyril S. Ku,Andrew T.Russo et al.“Performance Analysis of Digital Signal Processors Using SMV Benchmark”,International Journal of Signal Processing[J],Summer 2009
- [3] C.V.Goudar,Pankaj Rabha,Murali Deshpande et al.“SMV Lite:Reduced Complexity Selectable Mode Vocoder”,IEEE[J],Pro ICASSP-2006:PP.701-704
- [4] Selectable Mode Vocoder Service Option for Wideband Spread Spectrum Communication Systems[EB/OL].3GPP2,C.S003-0 V2.0. 2001
- [5] 日本无线工业及商贸联合会.ARIB STD-T64-C.S0030-0 V2.0[EB/OL]. 2001
- [6] 日本无线工业及商贸联合会.Software Distribution V3.5 for C.0030-0 V2.0 SMV Specification[EB/OL]. 2001

作者简介:

王小兵(1983-), 男, 湖北孝感, 硕士研究生, 主要研究方向: 语音信号处理;

黄 冰(1946-), 男, 江西樟树, 教授, 主要研究方向: 信号处理、计算机网络;

杜俊宇(1983-), 男, 山西五台, 硕士研究生, 主要研究方向: RFID 系统应用。

基于方向特征和局部walsh变换的虹膜识别新算法

王凤随¹ 李素文¹ 冯 桂²

(1 淮北煤炭师范学院 物理与电子信息学院, 安徽 淮北 235000;

2 华侨大学 信息科学与工程学院, 福建 泉州 362021)

摘 要: 虹膜纹理上的差异是不同虹膜相互区别的最主要原因和体现。因此, 要想获得很高的虹膜识别率, 就要有效地提取虹膜的纹理信息。通过对虹膜纹理单元的分析, 提出一种基于像素方向相关性和局部沃尔什变换提取虹膜纹理特征的新方法。在求取归一化虹膜图像中每个像素的局部沃尔什变换系数的基础上, 在纹理单元内估计变换系数的各阶矩能量测度作为虹膜的特征向量。实验结果表明, 算法能够有效地提高虹膜识别率, 可达 99%以上, 且算法简单、快速。

关键词: 虹膜识别; 局部沃尔什变换; 方向特征; 纹理分析; 特征提取; 识别率

A Novel Iris Recognition Algorithm Based on directional feature and Local Walsh Transform

Wang Feng sui¹ Li Su Wen¹ Feng Gui²

(1 Department of Physics and Electronics Information Huaibei Coal Industry Teachers' College, Huaibei Anhui, 235000, P.R.China; 2College of Information Science and Engineering, Huaqiao University, Quanzhou Fujian 362021, P.R.China)

Abstract: Irises can be distinguished mainly due to textural difference among them. In order to acquire a high recognition rate, the textural information of iris images must be extraction effectively. A novel iris texture feature extraction algorithm based on pixel's directional relativity and local walsh transform was presented according to analyze iris texture units. For each pixel in normalized iris images, the local walsh transform coefficients were computed, and their order moments were estimated as the iris's texture features in texture unit. According to experiment, the algorithm can enhance iris recognition rate, achieving above 99 percent. Furthermore, algorithm is simple and fast.

Keywords: iris recognition; local walsh transform; directional feature; texture analysis; feature extraction; recognition rate

1 引言

随着当今网络化进程的发展, 安全问题显得尤为重要, 基于生物特征识别的智能身份识别方法的研究正在受到人们的广泛关注。目前已经出现了不同的生物识别特征, 包括有声音、人脸、指纹、视网膜、签名、步态和虹膜等。作为重要的身份鉴别特征, 虹膜具有唯一性、

稳定性、可采集性和非侵犯性的优点,并且它的错误率是各种生物特征识别中最低的^[1]。因此,虹膜识别的研究已经得到学术界和企业界的高度重视。目前,国内外已有许多研究机构投入人力对虹膜识别进行研究,如剑桥大学的 Daugman 博士^[2]、美国 MIT 人工智能实验室的 Wildes 博士^[3]、澳大利亚 Queensland University 的 Boles^[4]、我国中科院自动化研究所的谭铁牛^[5]等。

对于虹膜特征提取和编码,存在两个主要的问题。一是哪些信息是虹膜纹理中最能区别于其它虹膜的特征,二是如何降低虹膜编码和匹配的算法复杂度,以达到实用性的目的。基于上述问题,本文针对虹膜图像纹理的分布特点,以纹理基元为单位,提出了基于像素相关性和局部沃尔什变换的虹膜纹理分析方法,具有运算简单、识别准确性高等优点。

2 虹膜图像预处理

虹膜图像预处理包括虹膜定位和归一化两个部分。如图 1(a)所示,是一幅采集而来的眼睛图,从图中不难看出,图像包含了一些我们不感兴趣的信息,如瞳孔,巩膜,眼睑等信息。为此,在特征提取之前,图像需要预处理以消除这些因素的影响。首先,应对虹膜的边界进行定位,即要找出瞳孔与虹膜之间(内边界)、虹膜与巩膜之间(外边界)的两个边界。通过采用 Canny 算子和 Hough 变换^[6]相结合的方法对虹膜图像定位。定位结果如图 1(b)所示。

由于光照强度及瞳孔膨胀的变化,虹膜的大小会发生变化,而且在虹膜纹理中发生的弹性变形也会影响虹膜模式匹配,必须对定位后的虹膜进行归一化,以补偿大小和瞳孔缩放所引起的变异。因为虹膜内外都为圆形,可以使用极坐标方法进行归一化。以瞳孔的中心为起点,与虹膜内外边界的交点坐标分别为 $(x_i(\theta), y_i(\theta))$ 和 $(x_o(\theta), y_o(\theta))$, 则利用下式:

$$\begin{cases} x(r, \theta) = (1-r)x_i(\theta) + rx_o(\theta) \\ y(r, \theta) = (1-r)y_i(\theta) + ry_o(\theta) \end{cases} \quad (1)$$

由式(1)可将虹膜图像中的每个点一一映射到极坐标 (r, θ) 中。结果虹膜区域展成如图所示的归一化矩形,使矩形的行对应虹膜的半径,列对应所取的角度。由于虹膜信息主要受到上眼睑的遮挡,为了更好的去除眼睑和睫毛等噪声信息,在归一化时,将角度定位在 $\theta \in [0, \pi/4]$ 和 $\theta \in [3\pi/4, 2\pi]$ 这个扇形区域内,如图 1(c)所示。归一化虹膜如图 1 所示,其中图 1(d)为未去除上眼睑的归一化图,图 1(e)是去掉上眼睑后的归一化结果。通过对比不难看出,通过选取一定角度,可以很好地去除虹膜上眼睑带来的大部分干扰信息。

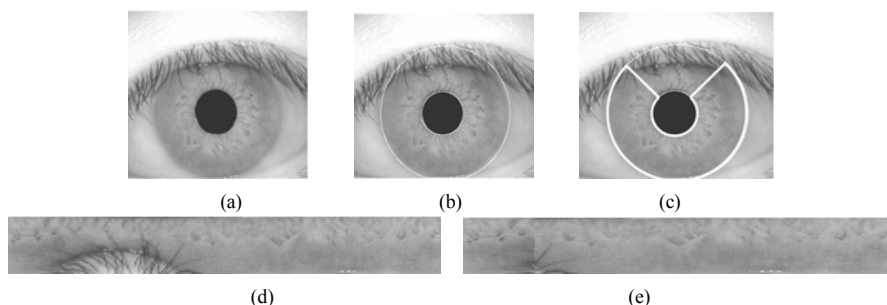


图 1 虹膜图像预处理

3 虹膜纹理特征提取与编码

3.1 方向特征与局部沃尔什变换

局部沃尔什变换充分考虑了像素间的相关性，进而提取出的特征能反映虹膜纹理的变化，属虹膜纹理方向特征，定义中考虑了虹膜在垂直、水平、45° 和 135° 等方向上的纹理变化。定义如下：局部沃尔什变换是定义在纹理单元^[7]上的一个 8 点的离散沃尔什变换，旨在提取纹理单元的局部纹理信息。

f_0	f_1	f_2
f_7	f_8	f_3
f_6	f_5	f_4

图 2 纹理单元中各个像素的标记方式

假设图像表示为 $\{f(x,y)|x=0,1,\dots,L-1,y=0,1,\dots,M-1\}$ ，其中 $f(x,y)$ 表示像素 (x,y) 的灰度值。按照图 2 所示的方式对纹理单元中各个像素进行编号，纹理单元中第 i 个像素的灰度值用 f_i 表示， $i=0,1,\dots,7$ 。则可以将纹理单元中各个像素的灰度值的空间关系属性转换成一维序列的形式。根据纹理单元中周围像素与中心像素的灰度差值可得到一个 8 点的一维序列：

$$f(n) = f_i - f_8, \quad i = 0,1,\dots,7 \tag{2}$$

对该序列进行离散沃尔什变换，变换点数为 $N=8$ ，变换系数为：

$$W(u) = \sum_{n=0}^7 f(n)Wal(u,n), \quad u = 0,1,\dots,7 \tag{3}$$

变换采用 Hadamard 编号的沃尔什函数作为基函数，则上式具有矩阵乘积的简单形式，

$$\begin{bmatrix} W(0) \\ W(1) \\ \vdots \\ W(7) \end{bmatrix} = H_8 \times \begin{bmatrix} f(0) \\ f(1) \\ \vdots \\ f(7) \end{bmatrix} \tag{4}$$

由于以上离散沃尔什变换是在纹理单元中进行的，故称之为局部沃尔什变换。因序列 $f(n)$ 反映了纹理单元中周围像素与中心像素的大小关系，体现了虹膜像素相关性特征。因此变换系数 $W(u)$ 反映了纹理单元中的像素值在列率域的变换特性，包含着图像的局部纹理信息。

3.2 特征提取和编码

纹理分析的关键是提取能够反映图像纹理属性的数量特征。实验表明，局部沃尔什变换系数并不能有效区分不同虹膜的纹理，然而不同纹理的局部沃尔什变换系数具有不同的统计特性，因此，通常的做法是用局部沃尔什变换系数的“能量测度”作为纹理特征。通过对部分图

基金项目：安徽高校省级自然科学基金计划项目 (KJ2008A114)；安徽省科学技术创新项目(2006KJ003TD)；福建省自然科学基金(A0610022)；

像的各个变换系数的各阶矩测试可知，变换系数的奇数阶矩并不具有很好的纹理，偶数阶矩则体现了不同虹膜图像的主要纹理差别，具有较较好的纹理鉴别能力，为此实验采用 2、4、6 阶矩作为虹膜的纹理特征。变换系数 $W(u|x,y)$ 的各阶中心矩在以像素 (x,y) 为中心的窗口中进行估计，窗口大小为 $W \times W$ ，变换系数 $W(u|x,y)$ 的 1 阶中心矩为：

$$B_1(u|x,y) = \frac{1}{W^2} \sum_{(i,j) \in W(x,y)} W(u|i,j) \quad (5)$$

k 阶中心矩为：

$$B_k(u|x,y) = \frac{1}{W^2} \sum_{(i,j) \in W(x,y)} [W(u|i,j) - B_1(u|x,y)]^k \quad (6)$$

其中， $W(x,y)$ 表示以 (x,y) 为中心的窗口，即小波系数的均值； $u=0,1,\dots,7$ 是局部沃尔什变换系数的序号，而 $k=2,3,\dots$ 为矩估计量的阶数。

基于上述思路，提出一种新的虹膜纹理特征提取的方法。具体算法如下：

Step1: 一维信号的形成。对归一化的虹膜图像，计算其周围的 8 个像素(标号方式见图 2)与中心像素的差值，并按照顺排列成一个 8 点长的一维序列。

Step2: 局部沃尔什变换。对 step1 的序列进行离散沃尔什变换，可得到对应的沃尔什变换系数，这些系数反映了虹膜图像在不同方向下的局部纹理信息。

Step3: 变换系数的能量测度。利用式(5)、(6)对各个纹理单元的变换系数分别计算各 2、4、6 阶矩能量测度。

Step4: 以 step3 总计算的各个纹理单元局部沃尔什变换系数的各阶矩作为虹膜的特征向量。

虹膜图像特征提取和编码的过程如图 3 所示：



图 3 特征提取和编码算法流程图

4 虹膜匹配

在获得了虹膜的特征向量之后，所要做的工作就是解决如何对虹膜进行分类的问题，即判断两个虹膜样本是否是来自同一个眼睛。把两个虹膜对应的特征向量作为特征点，那么可以用对应特征点之间的欧式距离来衡量两个虹膜的相似性。特征空间中两点间广义欧式距离为：

$$ED(V_A, V_B)(i) = \sqrt{(M_{iA} - M_{iB})^2 + (S_{iA} - S_{iB})^2} \quad (7)$$

考虑到虹膜各纹理单元纹理分布的不同，文中采用基于方差倒数的加权欧式距离分类器，定义如下：

$$WED(k) = \sum_{i=1}^N \frac{(f_i - f_i^k)^2}{(\delta_i^k)^2} \quad (8)$$

式中， f_i 是输入待测虹膜的第 i 个特征； f_i^k 是虹膜 k 的第 i 个特征； δ_i^k 是虹膜 k 的第 i 个特征的标准差； N 表示特征总数。

5 实验结果与讨论

实验是采用 CASIA 虹膜图像数据库，在 MATLAB7 下实现的。通过对 50 个人的眼睛、350 幅大小为 320×280 的虹膜图像组成的样本进行测试。实验共进行 12450 次比较，其中同一虹膜的前 3 幅用于生成虹膜特征值，后 4 幅用于识别，共进行 $50 \times 4 = 200$ 次比较，不同虹膜选择其中的 5 幅进行识别，共进行 $5 \times 49 \times 50 = 12250$ 次比较。

5.1 实验结果

在 CASIA 虹膜图像数据库中每一眼睛共由 7 幅不同虹膜图片组成，选取前 3 幅于第一阶段拍摄采集到虹膜图像作为虹膜特征数据库的训练集，后 4 幅第二阶段采集的图像则作为待识别的对象。对训练集中相同眼睛的 3 幅虹膜图片分别求出其特征值，然后对 3 幅虹膜图片相应的特征条带的特征值进行平均处理即可得到该虹膜的特征值。

为选择最佳阈值对虹膜进行识别，实验中分别计算出在不同的阈值下的错误拒绝率 FRR(false rejection rate)和错误接受率 FAR(false acceptance rate), 实验结果如图 4 所示。从图 4 可以看出算法的 FAR 随着 Hamming 距离阈值的增大而增加, 相反, FRR 随之减小。FAR 和 FRR 曲线的交叉点, 即等错率 EER(equal error rate)点很低, 在 1% 以下, 约为 0.35%。

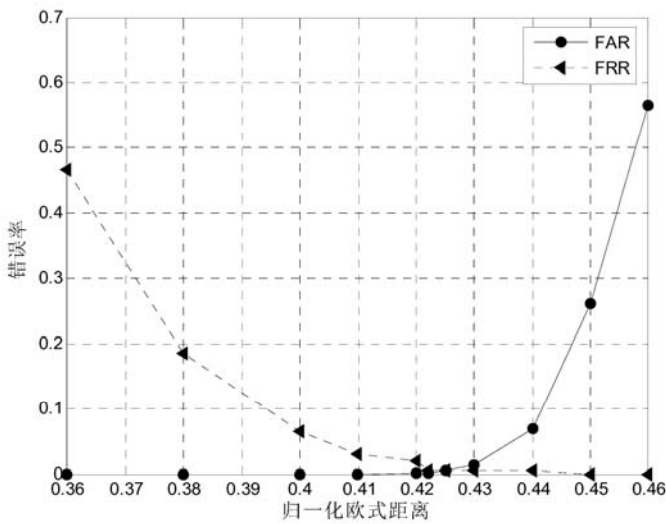


图 4 错误率曲线图

Pentium(R)4 3GHz, 512MB 的内存，操作系统为 Windows XP 的环境下测试算法的运行速度。其中特征提取的平均时间为 118ms，单次识别所需总时间约为 211ms，可满足实时识别的要求。

5.2 比较和讨论

为了测试和比较几种主要虹膜识别算法，中科院谭铁牛^[8]等人采集了一个虹膜识别数据库 (CASIA)，并且在这个图片库上测试了 Daugman, Boles 等人的算法。为了便于比较，表 1 列出了提出算法的测试结果及 Tan、Daugman、Boles 及 Wildes 的算法测试结果在正确识别率

CRR(correct recognition rate)和等错率 EER 上的对比。

表 1 几种算法的比较

方法	Daugman	Boles	Tan	Masek	提出算法
CRR (%)	100	92.64	99.43	98.86	99.65

通过比较可以看出，提出的算法明显优于 Boles 等算法，Daugman 的算法结果是目前最好的方法,同 Daugman 算法相比，虽然在准确率稍微有所降低，但提出的算法运行时间少，速度快，更容易满足实时性要求，易于实现。

通过分析发现，Daugman 通过利用一族多尺度的 Gabor 滤波器对虹膜图像进行局部相位编码来形成虹膜码。该方法达到了很高的识别率，但对图像的质量要求高。Boles 等人提出利用一维小波变换过零点检测的方法，此方法有效地克服了亮度变换的干扰且具有一定的鲁棒性，但又存在匹配精度不高，表达式无法体现虹膜纹理方向性和对纹理灰度变化敏感等缺点。Wildes 等人利用四层拉普拉斯金字塔分解的方法，以塔式分解系数作为纹理特征，方法克服了灰度的影响，但运算量大，匹配算法复杂，不利于实际应用。

6 结论

实验结果表明，提出的基于局部沃尔什变换的方法从速度和识别准确性上都可以满足识别的要求，而且，由于计算局部沃尔什变换系数只需要进行灰度值的加减运算，因此算法具有计算量小，特征提取速度快，计算简单的优点。算法在提取纹理单元的局部纹理信息时，利用了周围像素与中心像素的灰度差值，充分考虑了虹膜纹理的方向性和纹理的变化；同时利用局部沃尔什变换，将局部纹理信息从空间域变换到列率域进行分析，为虹膜特征提取的方法上提供了一个新的思路。

参 考 文 献

[1] JAIN A.Biometrics personal identification in networked society[M]. Kluwer Academic Publishers , 1999.

[2] J. Daugman. Statistical Richness of Visual Phase Information: Update on Recognizing Persons by Iris Patterns [J]. Computer Vision, 2001, vol. 45: 25-38.

[3] Wildes,P. Iris Recognition: An Emerging Biometric Technology [J]. PROCEEDINGS OF THE IEEE,1997,85 (9):1348-1363.

[4] W. Boles, B. Boashash. A human identification technique using images of the iris and wavelet transform [J]. IEEE Trans. Signal Processing, 1998, 46 (4): 1185-1188.

[5] L. Ma, Y. Wang, T. Tan, “Iris Recognition Using Circular Symmetric Filters”, Proc. of the 16th International Conference on Pattern Recognition, Vol. II, pp.414-417, 2002.

[6] Qichuan Tian, Quan Pan, Yongmei Cheng et al. Fast algorithm and application of Hough transform in iris segmentation [J]. Machine Learning and Cybernetics, 2004. Proceedings of 2004 International Conference on Vol. 7: 3977-3980.

- [7] Hui Yu, Mingjing Li, Hongjiang Zhang, et al., "Color Texture Moments for Content Based Image Retrieval", Proc of the IEEE International Conference on Image Processing. Rochester, USA, pp. 929-932, 2002.
- [8] L .Ma, T .Tan, et al. Efficient Iris Recognition by Characterizing Key Local Variations [J]. IEEE Trans. on Image Processing, 2004, Vol. 13: 739-750.

作者简介:

王风随(1981—), 男, 安徽省宿州市人, 硕士, 讲师, 主要研究方向: 图像处理与模式识别。E-mail: fengsuw@163.com;

李素文(1968—), 女, 博士, 教授, 主要研究方向: 光电信息处理;

冯 桂(1960—), 女, 博士, 教授, 硕士生导师, 主要研究方向: 图像处理, 模式识别, 信息安全等。

变长小生境算法用于图像超分辨率复原

张煜东 吴乐南 奚 吉 王水花

(东南大学信息科学与工程学院 南京 210096)

摘 要: 为了更好地实现图像超分辨率复原, 本文提出一种新的算法如下: 首先, 证实了采用差值编码优于直接编码, 并且利用变长编码可以节约 25% 的存储空间; 其次, 将变长染色体算法与小生境算法相结合, 提出一种变长小生境算法; 最后, 为了进一步保证收敛性和快速性, 采用踏脚石模型的伪并行化方式。对单帧图像复原与多帧超分辨率复原的实验表明, 本文算法优于传统的遗传算法、变长染色体算法、小生境算法。而且, 踏脚石模型有效, 不仅能增加算法成功率, 而且能加快收敛。

关键词: 超分辨率复原; 遗传算法; 变长染色体算法; 小生境算法; 踏脚石模型

The messy-niche algorithm used for image super-resolution restoration

Zhang Yu dong Wu Le nan Xi Ji Wang Shui hua

(School of Information Science & Engineering, Southeast University, Nanjing 210096, China)

Abstract: In order to realize image super-resolution more efficiently, a novel hybrid algorithm was proposed in this paper—first, the difference coding was verified to be more excellent than direct coding, and the chromosome of variant length can save 25% restoration room—next, the messy algorithm was combined with niche algorithm to produce a novel method, namely, the messy niche algorithm—finally, a stepping-stone model of pseudo-parallel computation was utilized in interest of enhancing the convergence and fastness. Experiments on single-frame image restoration and multi-frame image super-resolution restoration showcased that this proposed algorithm was superior to genetic algorithm, messy algorithm, and niche algorithm. Moreover, the stepping-model was effective, and it was capable of not only improving the successful rate but also fastening the convergence.

Keyword: super-resolution restoration; genetic algorithm; messy algorithm; niche algorithm; stepping-stone model

1 引言

图像超分辨率复原指采用一幅或多幅低分辨率图像来获得原始高分辨率图像的过程。若只采用一幅低分辨率图像, 则模型退化为单帧图像复原。

1974 年, Besag 通过 Markov 随机场和 Gibbs 随机场理论, 指出在最大后验概率意义下,

图像超分辨率等价于一个极小化问题。随后,采用条件迭代模型(ICM)来求解。然而问题不是凸的,易陷入局部最优;另外复原质量受噪声影响很大,因此传统的求解方法不适用。

随着智能优化这一领域的活跃,由于其在寻找全局最优解方面的突出表现,越来越多的学者将智能优化算法应用于图像复原。Chen等提出用GA进行灰度图像的复原⁰。Zhao提出采用遗传算法优化堆栈滤波器(Stack filter),然后将最优堆栈滤波器用于图像复原⁰。Kim等提出用GA设计等照度线约束下的区域移除和复原算法⁰。Isgro提出用分布式ad-hoc遗传算法复原被竖线划伤的图像⁰。

为了进一步加快收敛,且保证能找到全局最优。提出一种新的思路:采用变长染色体来编码,可节约空间、计算量;另外引进小生境算法,来增强算法的收敛能力;最后,采用一种踏脚石信息交换模型,使算法伪并行运行。

2 背景介绍

2.1 图像退化模型

实际中,低分辨率(LR)图像是由同一场景的 HR 图像经多种退化作用后的结果,而图像超分辨率复原也就是要明确这些退化的原因,再由 LR 复原出 HR。如图 1,HR 图像 X 经过几何形变 G 、模糊 B 和降采样 D 后,再混入加性零均值高斯白噪声 N ,得到 LR 图像 Y ,有

$$Y = GBDX + N \tag{1}$$

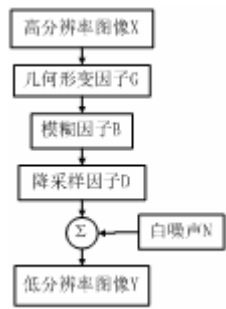


图 1 图像超分辨率复原的退化模型

假设几何形变 G 已经校正,并记降质因子 $H=BD$ 。再假设对于每幅 LR 图像 Y_i ,均由一幅 HR 图像 X 经过不同类型的降质因子 H_i 与噪声 N_i 形成,有

$$Y_i = H_i X + N_i \tag{2}$$

因此,可如下定义适应度函数来求原始 HR 图像:

$$E(\hat{X}) = \sum \|N_i\|^2 = \sum \|Y_i - H_i \hat{X}\|^2 \tag{3}$$

一般添加一个正则化项,来抑制图像噪声,因此适应度函数改写为:

$$E(\hat{X}) = \sum \|Y_i - H_i \hat{X}\|^2 + \lambda \|L\hat{X}\|^2 \tag{4}$$

注意,若 LR 图像仅有 1 幅,则图像超分辨率复原退化为单帧图像复原,此时适应度函数为:

$$E(\hat{X}) = \|Y - H\hat{X}\|^2 + \lambda \|L\hat{X}\|^2 \tag{5}$$

式(4)与(5)中, \hat{x} 是HR的估计, L 为正则化算子, 通常为Laplacian算子以使图像平滑。 λ 是正则化参数。显然, 适应度函数 E 越小, 恢复性能越好。

2.2 变长染色体算法

生物进化过程中, 其染色体的长度并不非固定不变, 而是随着进化过程慢慢变化着。传统的遗传算法(GA)认为每个染色体长度是固定的, 虽然给解决问题带来简化, 但可能会造成存储量和运算效率的浪费。在有些场合问题本身的性质决定了染色体长度是变化的, 这时采用 Messy 算法不但能节约存储空间, 且能充分提高计算效率。

Messy算法⁰中, 保留简单遗传算法(SGA)中的选择算子和变异算子, 舍弃交叉算子, 引进如图 2 所示的新算子: 切断算子与拼接算子, 作为产生新个体的主要遗传算子。

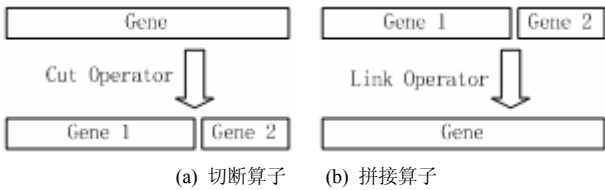


图 2 MA 算法中的新算子

切断算子以某一预先指定的概率, 在变长染色体中随机选择一个基因, 在该处将个体的基因型切断, 使之成为二个个体的基因型。拼接算子以某一预先指定的概率, 将两个个体的基因型连接在一起, 使它们合并为一个个体的基因型。

通过切断算子和拼接算子, 使得染色体长度可变, 如果变化后的适应度函数较优, 则这种变化就会固定下来, 然后 MA 算法会继续变化整个群体的分布函数, 最后找到全局极小。

2.3 小生境算法

本文提出一种小生境算法基于以下思想: 首先比较个体间距, 若在预先指定的惩罚距离 L 之内, 再比较两者之间的适应度大小, 并对其中适应度较低的个体施加一个较强的罚函数, 使其适应度极大地降低。这样, 在距离 L 之内将只存在一个优良的个体, 从而既维护了群体的多样性, 又使得各个体之间保持一定的距离, 并使得个体能够在整个约束空间分散开来。

3 新的编码方式

3.1 差值编码

理想情况下, 最终进化算法搜索 \hat{x} 直至找到原始清晰图像 X , Y 是已知的 LR 图像, 令 $Z=X-Y$ 。显然 Z 与 X 可以互相转化, 因此直接对 X 编码搜索 (称为直接搜索), 和对 Z 编码搜索, 得到结果后再计算 X (称为差值搜索), 两者结果一样, 如图 3 所示。

然而, X 的分布一般较为平坦, Z 的分布一般较为集中, 如图 4 所示。

图 4 (a)表示原始 HR 图像 X , 图 4 (b)表示 X 被 3×3 窗口方差为 0.5 的高斯模糊算子模糊, 并加上均值为 0 方差为 0.01 的高斯白噪声, 最后得到的 LR 图像 Y , 图 4 (c)表示两者之差 $Z=X-Y$ 。

图 4 (d)与(e)分别表示 X 与 Z 的直方图, 可见, Z 比 X 更加集中, 且近似高斯分布, 根据遗传算法搜索理论, 对 Z 的搜索比对 X 的搜索收敛更快。

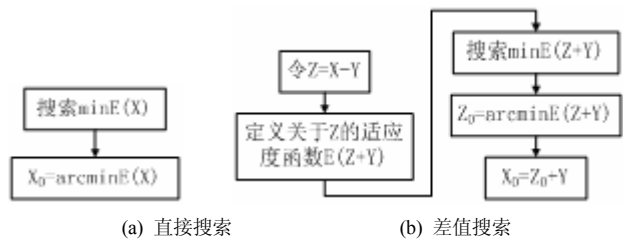


图 3 两种搜索方式

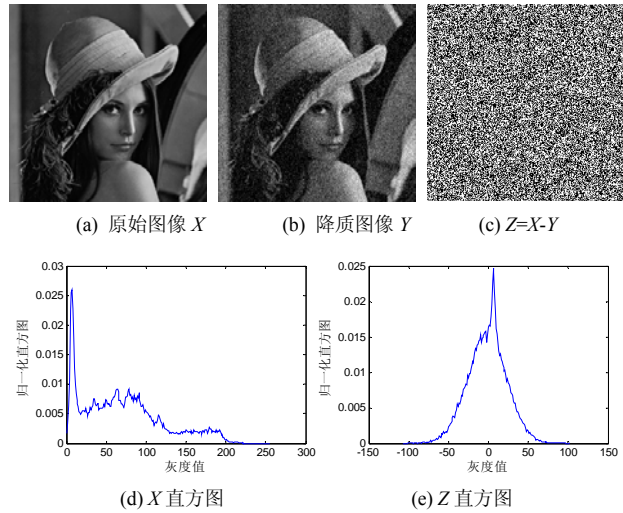


图 4 X 与 Z 的比较

3.2 差值变长编码

上文阐述了差值搜索优于直接搜索的原因, 本节阐述采用变长编码优于固定长度编码的原因。

如图 4(e)所示, 如果忽略 $[-128, 127]$ 区间之外的直方图, 则使用固定长度的编码方式, 须取 8 位才能保证信息不丢失。但此时大部分差值集中在 0 附近, 因此会造成存储空间冗余, 且增加了运行开支。

利用信息论的变长编码理论, 计算上百幅图像的差值信息, 发现对差值 Z 而言, 平均码长仅为 6.09 位。实际使用中, 发现采用 6 位的信息一般已经足够, 并规定染色体拼接后的上限为 8 位。因此, 可节约 25%的存储空间, 同时减少运算量。

4 变长小生境算法

4.1 算法流程

将 MA 与 NA 的思想相结合, 即得到本文所提出的变长小生境算法(MNA)的思想。具体

框架如图 5 所示。

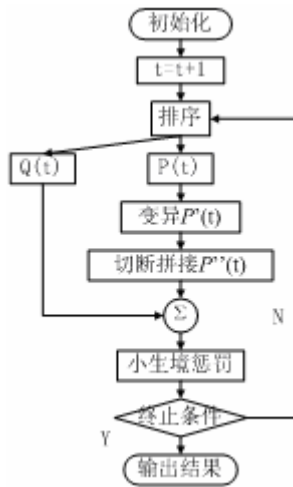


图 5 MNA 算法框架

算法 1 (MNA):

Step 1 初始化。设置进化代数计数器 $t=0$ ；随机生成 $M+N$ ($N < M$) 个初始个体组成初始群体 $P(t)$ ，长度全部为 $k=6$ 。求出各个个体的适应度 F_i ($i=1, 2, \dots, M$)；

Step 2 进化代数递增。 $t=t+1$ ；

Step 3 排序。依据各个个体的适应度对其进行降序排序，记忆前 N 个最优个体，记为 $Q(t)$ ；其余 M 个非最优个体，记为 $P(t)$ ；

Step 4 变异运算。对 $P(t)$ 进行变异运算，得到 $P'(t)$ ；

Step 5 切断、拼接运算。对群体 $P'(t)$ 施加切断算子和拼接算子，得到 $P''(t)$ ，其中限制染色体拼接后上限为 8 位；

Step 6 小生境惩罚运算。将 Step 5 得到的 M 个个体 $P''(t)$ 与 Step 2 所记忆的 N 个个体 $Q(t)$ 合并，得到一个含有 $M+N$ 个个体的新群体；对这 $M+N$ 个个体，求出每两个个体 X_i 和 X_j 的海明距离：

$$\|X_i - X_j\| = \sqrt{\sum_{k=1}^l (x_{ik} - x_{jk})^2} \quad (6)$$

当 $\|X_i - X_j\| < L$ 时，比较个体 X_i 和 X_j 的适应度，并对其中适应度较低者处以罚函数：

$$\begin{aligned} F(X_i) &= \text{Penalty if } F(X_i) < F(X_j) \\ F(X_j) &= \text{Penalty if } F(X_j) < F(X_i) \end{aligned} \quad (7)$$

Step 7 终止条件判断。若不满足终止条件，则转 Step 2；若满足终止条件，则输出计算结果，算法结束。

4.2 收敛性证明

算法中引进了最优 N 保护策略，具体可见 Step 2。下述定理指出若不采用最优 N 保护策略，则无法保证算法进化到全局极小；若采用最优 N 保护策略，则算法进化到全局极小的概率为 1。

定理 1: 若不使用最优保护策略, 则算法进化到全局极小的概率小于 1。

证明: 将每一代群体看作一种状态, 则可以把进化过程作为一个齐次 Markov 链 $\{P(t), t \geq 0\}$, $P(t)$ 的各种状态 I 可以划分为: 包含全局极小的状态 I_o 和不包含全局极小的状态 I_n 。

进化过程中, 令选择、变异、切断、拼接和小生境算子分别为矩阵 $S=\{s_{ij}\}$, $M=\{m_{ij}\}$, $C=\{c_{ij}\}$, $A=\{a_{ij}\}$ 和 $N=\{n_{ij}\}$, 显然上述矩阵都是随机矩阵, 即满足每行之和等于 1。算法每进化一代的变换矩阵即为 $R=SMCAN=\{r_{ij}\}$ 。

由于 $m_{ij}=p_m^{H(i,j)}(1-p_m)^{1-H(i,j)}>0$, 所以变异矩阵的元素恒正。式中 p_m 为单个染色体的变异概率, $H(i,j)$ 为状态 i 与状态 j 的海明距离。所以 $r_{ij}=\sum_k \sum_l \sum_m \sum_n s_{ik} m_{kl} c_{lm} a_{mn} n_{nj} > 0$,

又 $P(t)$ 稳态时满足稳态方程: $P_f(\infty)=\sum_{i \in I} P_i(\infty) r_{ij}$, 由于 $r_{ij}>0$, 所以 $P_j(\infty)>0$ 。当 $j \in I_n$ 时,

有 $\lim_{t \rightarrow \infty} \{P_t \in I_n\} > 0$, 从而 $\lim_{t \rightarrow \infty} \{P_t \in I_o\} < 1$ 。证毕。

定理 2: 若使用最优 N 保护策略, 则算法进化到全局极小的概率等于 1。

证明: 根据最优保护的定义, 随机过程的状态转移概率满足:

$$\begin{aligned} r_{ij} &> 0 \quad (\forall i \in I, \forall j \in I_o) \\ r_{ij} &= 0 \quad (\forall i \in I_o, \forall j \in I_n) \end{aligned} \quad (8)$$

于是有 $\lim_{t \rightarrow \infty} r'_{ij} = 0 (\forall i \in I, \forall j \in I_n)$, 式中 r'_{ij} 表示 t 步从状态 i 转移到状态 j 的概率。所以

$\lim_{t \rightarrow \infty} P_f(t) = 0 (j \in I_n)$, 即 $\lim_{t \rightarrow \infty} \{P_t \in I_o\} = 1$, 证毕。

5 基于踏脚石模型的伪并行化

为提高遗传算法速度, 开发了许多不同的并行遗传算法, 分为标准型并行方法和分解型并行方法两类。SPA 不改变 SGA 的基本结构特点, 在一个总体环境中实现进化, 使用一个统一的群体, 实现时需要一个全局存储器, 还需要一个统一的控制机构来协调群体的进化过程和通讯过程, 出发点是从局部的角度开发个体进化的并行性。

DPA 将整个群体划分为几个子群, 各个子群分配在独立的处理机(称为并行算法)或同一台计算机的不同线程内(称为伪并行算法)上进行各自进化操作, 出发点是从全局角度开发群体进化的并行性。由于保持了各处理机上子群体进化的局部特性, 且各个子群存在交互, 使得算法一方面能提升寻找到全局最优的概率(成功率), 另一方面能大幅降低搜索时间。

经典方法有: 图 6 所示的踏脚石群体模型(Stepping-stone model)⁰、岛屿群体模型(Island Model)和邻居群体模型(Neighborhood Model)。为了测试不同模型的优劣, 对标准测试函数 Six-hump Camel Back 进行寻优, 得到表 3 结果。

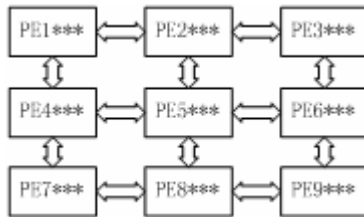


图 6 踏脚石模型(其中***号表示任意个数的任意个体)

表 3 采用并行模型后算法性能提升比较（100 次试算）

算法	收敛性能		时间性能	
	早熟次数	成功率	平均代数	运行时间(s)
SGA 串行模型	40	60%	51	0.0581
岛屿并行模型	3	97%	39	0.0248
踏脚石并行模型	5	95%	35	0.0232
邻居并行模型	12	88%	4	0.0027

可以看出，抗早熟能力最强的是岛屿模型，但运算最慢；运算最快的是邻居模型，但抗早熟能力差。而踏脚石模型兼顾了两者的优点，既保证了较强的抗早熟能力，又有较快的运算速度⁰。另外，这 3 种并行模型的成功率与运行时间均优于传统的串行模型。

本文在单机条件下，在软件环境上借鉴踏脚石模型作为并行计算方法，开发出一种基于踏脚石模型的伪并行(Stepping-Stone Pseudo-Parallel, SSPP)算法，使其既继承了踏脚石并行模型的优点，又省去了实际并行机的昂贵开销。

算法 2（MNA-SSPP 算法）：

- Step 1 遗传代数计算器初始化： $t=0$;
- Step 2 将群体 $R(t)$ 划分为子群体： $R(t)=\{R_1(t), R_2(t), \dots, R_n(t)\}$ ，其中 n 为分组数;
- Step 3 各自进化。每组算法均采用 MNA 计算，参数设置相同;
- Step 4 交换信息。若 t 满足交换代数，则 $R(t)=\text{exchange}[R(t)]$;
- Step 5 终止条件判断：若不满足则 $t=t+1$ ，转向 Step 3；若满足则输出优化结果，算法结束。

Step 4 中信息交换的对象、内容、频率和信息量均由不同的踏脚石模型预先定义。本文根据实验结果，最终确定参数如下选取：对邻近的计算机，内容为从最优的几个个体中任选一个进行交换，频率为每五代交换一次。

6 实验

6.1 参数设置

采用本文算法（算法 2）进行超分辨率复原，对比算法采用GA⁰、MA⁰、NA。参数设置见表 4。注意由于算法采用最优N保护策略，所以“90+10”指将 100 个个体分为两组，其中最优群体大小N=10；MA的变异率 0.9 是对整个群体而言，由 90/100=0.9 得来，即非最优群体大小M=90 全部变异。

表 4 各算法的参数设置

算法	群体大小	进化代数	交叉率	变异率	切断率	拼接率	惩罚距离	子群数	信息交换间隔
GA	100	1000	0.8	0.01	-	-	-	-	-
MA	100	1000	-	0.9	0.5	0.5	-	-	-
NA	100	1000	0.8	0.01	-	-	20	-	-
算法 1	90+10	1000	-	0.9	0.5	0.5	20	-	-
算法 2	90+10	1000	-	0.9	0.5	0.5	20	5	5

终止条件设置为下列 2 个条件之一满足即可：1) 迭代次数达到或超过给定的最大迭代次数；2) 在超过 10 次迭代中，群体的适应度函数改善量小于 10^{-4} 。

实验中为了加速处理，将图像划分为 16×16 的子图，相邻子图有 2 像素的重叠，以抵消边界效应。另外，考虑到进化算法本身的随机性，每个算法运行 10 次，取平均值作为结果。

6.2 单帧复原

6.2.1 Mar 图像

采用标准 Mar 图像，大小为 128×128 ，取方差为 0.5 的 3×3 高斯模糊核，加均值 0、标准差 0.001 的高斯白噪声，正则化参数 $\lambda=0.005$ ，评价参数选取峰值信噪比(PSNR)。

复原结果示于图 7，PSNR 与运行时间示于表 5。主观上看本文算法的复原结果优于 GA、MA、NA；客观上看，本文算法的 PSNR 最高。同时算法 1 时间耗费太长，而算法 2 的时间仅有算法 1 的 2/3 左右，因此基于踏脚石模型的伪并行算法有效。

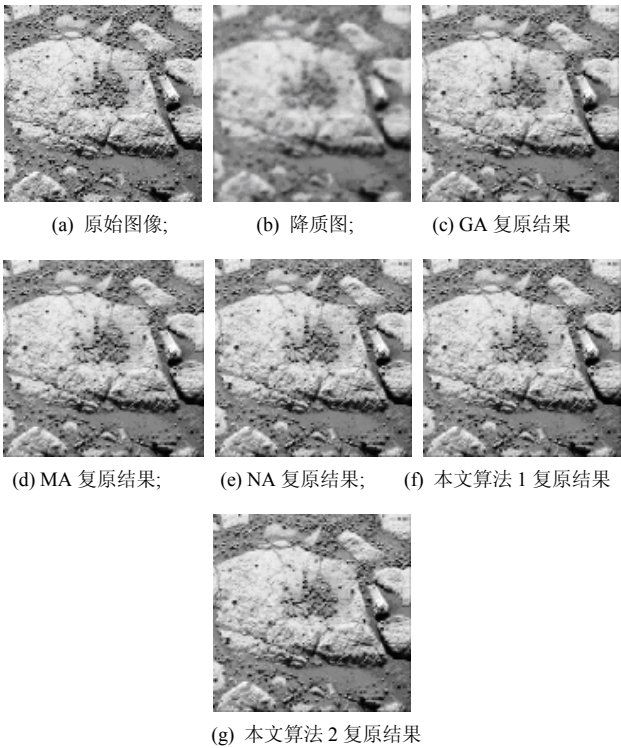


图 7 Mar 图像复原

表 5 Mar 各算法比较

算法	GA	MA	NA	算法 1	算法 2
PSNR(dB)	24.0513	25.8842	25.087	26.3313	26.5924
运行时间(s)	775	826	278	889	682

6.2.2 Flower图像

再以 Flower 图像作为测试，降质参数设置与算法的参数设置皆与上例相同。复原结果见图 8，比较结果见表 6。再次可见，本文提出的 2 个算法 PSNR 均最高，且采用伪并行计算后，节约了一半左右的时间，令人满意。

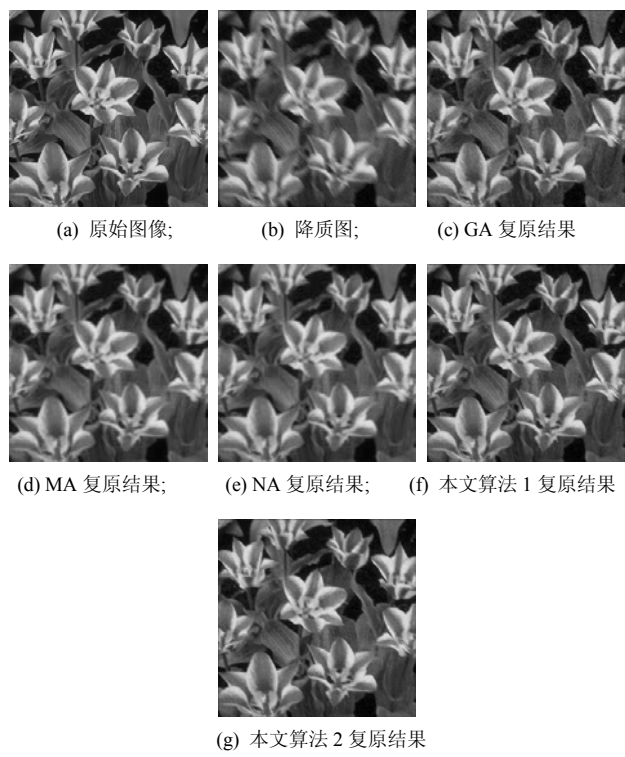


图 8 Flower 复原结果

表 6 Flower 各算法比较

算法	GA	MA	NA	算法 1	算法 2
PSNR(dB)	25.365	25.642	25.770	26.224	26.317
运行时间(s)	483	783	263	989	534

6.2.3 多帧超分辨率复原

采用 128×128 的Papav图像，模糊核选取长度在[3, 7]内任意变化，角度在[-180, 180]内任意变化的运动模糊函数，噪声选择均值 0 标准差 0.001 的高斯白噪声。10 幅降质图像中每帧的模糊函数与噪声均不同，但是假设模糊函数已知。关于如何从降质图像中提取模糊核的方法，请参见文献 0。

从 20 幅降质图像中任选 2 幅示于图 9。由于采用了多帧超分辨率图像复原，因此复原效果总体优于单帧图像复原，见图 10 与表 7。同时，我们也发现本文算法与其他算法的差距明显拉开，优于其余算法 4-5dB。这是因为当运算量增大时，本文算法的有效性与快速性得以体现。

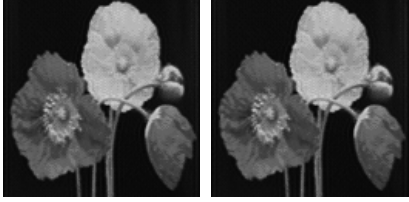


(a) 原图 (b) 模糊函数长度为 7 方向 45 (c) 模糊函数长度 5 方向-45

图 9 多帧 Papav 降质图像



(a) GA 复原 (b) MA 复原 (c) NA 复原



(d) 本文算法 1 复原 (e) 本文算法 2 复原

图 10 Papav 图像超分辨率复原结果

表 7 Papav 图像超分辨率复原各算法比较

算法	GA	MA	NA	算法 1	算法 2
PSNR(dB)	31.264	32.486	32.689	36.592	37.087
运行时间(s)	1036	1862	799	3145	1077

7 结论

本文针对图像超分辨率复原的正则化模型，提出一种用于求解的变长小生境算法。为了进一步保证收敛性与快速性，采用踏脚石模型实现伪并行运算。对单帧图像复原与多帧超分辨率复原的实验证实，算法 2 的复原效果最佳，且运行时间相对较短。

参 考 文 献

[1] Chen Y W, Nakao Z, Arakaki, et al. Restoration of gray images based on a genetic algorithm with Laplacian constraint [J]. Fuzzy Sets and Systems, 1999, 103(2): 285-293.

[2] Zhao C H, Zhang W C. Using genetic algorithm optimizing stack filters based on MMSE criterion [J]. Image and Vision Computing. 2005, 23(10): 853-860

[3] Kim J B, Kim H J. Region removal and restoration using a genetic algorithm with isophote constraint [J]. Pattern Recognition Letters, 2003, 24(9): 1303-1316

[4] Isgro F, Tegolo. A distributed genetic algorithm for restoration of vertical line scratches [J]. Parallel Computing.

- [5] Day R O, Lamont G B, Extended multi-objective fast messy genetic algorithm solving deception problems [J]. Evolutionary Multi-Criterion Optimization. 2005: 96-130
- [6] Aoki K, Nakahashi W. Evolution of learning in subdivided populations that occupy environmentally heterogeneous sites [J]. Theoretical Population Biology. 2008, 74(4): 356-368
- [7] Benjamin H, Kevin W. Synapsing variable-length crossover: Meaningful crossover for variable-length genomes [J]. IEEE Transactions on Evolutionary Computation. 2007, 11(1): 118-131
- [8] Shin H C, Prager R, Ng J. Sensitivity to point-spread function parameters in medical ultrasound image deconvolution [J]. Ultrasonics. 2009, 49(3): 344-357

一种基于形状信息的前景检测方法

刘乐元 桑 农

(华中科技大学 图像识别与人工智能研究所, 湖北, 武汉, 430074)

摘 要: 提出了一种基于形状信息的前景检测方法. 在对场景进行边缘检测的基础上, 以由形状上下文 (Shape Context) 描述的局部形状信息作为特征建立背景模型, 并用背景消减法 (background subtraction) 来检测前景. 与高斯混合模型 (Gaussian Mixture Model, GMM) 和局部二进制模式 (Local Binary Pattern, LBP) 的比较实验表明, 该方法具有较好的鲁棒性, 精确性和实时性.

关键词: 模式识别与智能系统; 前景检测; 背景消减; 形状上下文

A Shape-Based Method for Foreground Detection

Liu Le Yuan Sang Nong

(Institute for Pattern Recognition and Artificial Intelligence, Huazhong University of Science and Technology, Wuhan, 430074, China)

Abstract. A novel and efficient shape-based method for background modeling and foreground detection is presented in this paper. Scene background is modeled by local shape structure which is represented by shape context histograms, based on edge detection results. Foreground in form of edges is detected by comparing the shape context histogram of each edge point in the new frame against the background models. Comparisons using both visual and numerical evaluation demonstrate that our approach outperforms Gaussian mixture model (GMM) and Local Binary Pattern (LBP). Furthermore, the proposed method is capable of real-time processing.

Keywords: pattern recognition and intelligence system, foreground detection, background subtraction, shape context

1 引言

前景检测是智能视频监控系统的—个至关重要的模块, 其输出结果将作为后续处理 (如目标识别、跟踪) 的输入, 因此前景检测模块直接影响着后续处理甚至整个系统的性能. 对于使用固定摄像头的系统而言, 背景消减 (background subtraction) 是解决前景检测问题的—个有效技术手段. 背景消减的关键是建立有效的背景统计模型, 并实时维护和更新模型以适应变化的环境. 根据建立背景模型时所使用的信息, 现有方法可以粗略地分为基于单像素的方法和基于邻域的方法.

基于单像素的方法在背景建模时仅以单个像素的时间序列观测值作为样本进行统计学习.

理想情况下最简单的单像素模型可以用单个的高斯分布来表示^[0],但是,当场景中包含大量杂乱的微小运动(如树叶摆动等)的时候,单高斯模型难以准确描述像素的实际状态.为应付这种情况,Staussfer等人提出了以多个高斯函数加权的形式来表示像素观测值的概率分布情况的高斯混合模型(Gaussian Mixture Model ,GMM)^[0]. GMM自提出之后得到了广泛的应用,并有大量的文献围绕改进GMM算法进行了研究^[0,0].为应付更为杂乱的场景,文献[0,0]中提出了用核密度估计(Kernel Density Estimation, KDE)来学习背景模型的方法.KDE属于非参数估计方法,能更准确地反映像素状态变化的历史过程,但运算量以及所需的存储空间均远远大于GMM.除了像素灰度值和颜色外,像素的空间梯度^[0]、边缘^[0]等特征也被用于基于单像素的背景建模方法中。

基于邻域的方法在提取描述背景像素统计特性的特征时考虑其空间/时间邻域像素的信息.在文献[0]中,场景被分为若干等大小的小方块,并通过对比各个小方块内的边缘直方图来检测前景.Heikkilä^[0]等利用局部二进制模式(Local Binary Pattern, LBP)作为特征来建立背景模型.由于LBP算子具有抗光照变化等良好特性,该方法取得了的效果.此外,还有一些方法^[0,0]在对一个背景像素进行建模时同时考虑其时间和空间邻域的信息.因为使用了更多的信息,基于邻域的方法比基于单像素的方法具有更好的鲁棒性.但基于邻域的方法通常计算量大,限制了此类方法在诸如智能视频监控等实时系统中的应用。

本文首次选用形状作为特征来进行背景建模和前景检测.对背景而言,形状是一个非常稳定的特征,受光照变化等的影响较小;对前景来说,形状具有较强的区分性,例如,与背景具有相似形状的前景出现的概率比与背景具有相似颜色的前景出现概率要小得多.实验表明,本文方法具有较好的鲁棒性,精确性和实时性。

2 基于形状信息的前景检测方法

图1给出了本文方法的流程.首先利用SUSAN算法^[0]进行边缘检测,在此基础上用形状上下文(Shape Context)^[0]直方图来描述背景的局部形状特征.背景模型由一组形状上下文直方图组成.通过当前帧的形状上下文直方图与背景模型直方图的匹配来进行前景/背景分类,并根据分类结果更新背景模型.得到的以边缘形式表达的前景可以直接供基于边缘表达(edge-based)的后续处理使用.但对于那些基于区域表达(region-based)的后续处理来说,则需转换为以区域或块(blob)形式的表达。

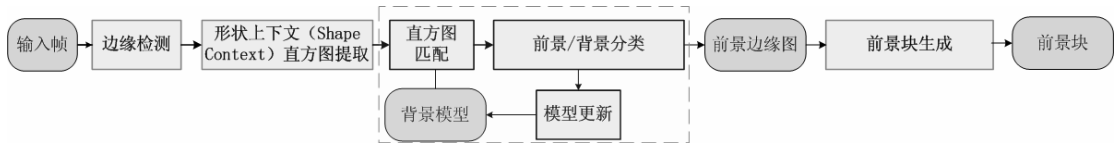


图1 基于形状信息的前景检测算法流程

2.1 形状上下文

形状上下文 (Shape Context)^[0]是一种广泛用于图匹配和目标识别的形状描述子.形状中的某采样点 p 的形状上下文直方图描述的是其邻域(8p)内其它采样点在极坐标系中的相对位置分布情况:

$$h_p(k) = \#(q : q \in \partial p, q \neq p(q-p) \in \text{bin}(k)) \quad (1)$$

对于形状上的两个点 p 和 q 的形状上下文直方图, 通常 χ^2 距离来度量它们的相似性:

$$D_{\chi}(\bar{h}_p, \bar{h}_q) = \frac{1}{2} \sum_{k=1}^E \left(\frac{[h_p(k) - h_q(k)]^2}{[h_p(k) + h_q(k)]} \right) \quad (2)$$

传统的形状上下文有一个缺点, 即当形状上的点刚好位于bin的 (方向或半径) 边缘的时候, 会使得原本很相似的两个形状计算得到的形状上下文直方图的差异却很大(如图 2(a)所示). 为解决这个问题, 文献[0]中给出了一种解决方案: 计算直方图的时候让位于角度bin边缘附近的采样点同时往两个相邻的bin投票(如图 2(b)). 但这个方法并没有从根本上解决问题. 为此, 本文采用二维线性插值的方法来计算形状上下文直方图:

$$\begin{cases} h(r_1, \theta_1) = h(r_1, \theta_1) + \frac{r_2 - r_1}{r_2 - r_1} \cdot \frac{\theta_2 - \theta}{\theta_2 - \theta_1} \\ h(r_1, \theta_2) = h(r_1, \theta_2) + \frac{r_2 - r_1}{r_2 - r_1} \cdot \frac{\theta - \theta_1}{\theta_2 - \theta_1} \\ h(r_2, \theta_1) = h(r_2, \theta_1) + \frac{r - r_1}{r_2 - r_1} \cdot \frac{\theta_2 - \theta}{\theta_2 - \theta_1} \\ h(r_2, \theta_2) = h(r_2, \theta_2) + \frac{r - r_1}{r_2 - r_1} \cdot \frac{\theta - \theta_1}{\theta_2 - \theta_1} \end{cases}, \quad \begin{matrix} \theta_1 \leq \theta \leq \theta_2 \\ r_1 \leq r \leq r_2 \end{matrix} \quad (3)$$

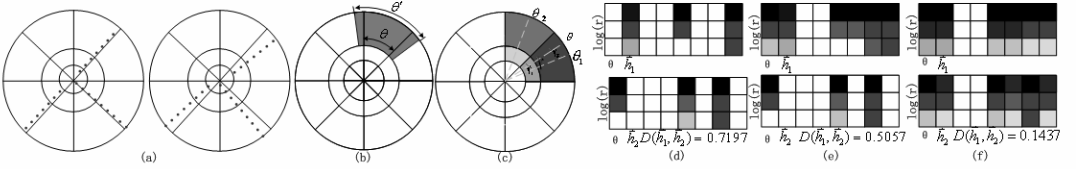


图 2 (a) 两个相似的形状 (b) 文献[16]中改进的形状上下文 (c) 本文用二维线性插值改进的形状上下文 (d) 图(a)中两个形状用原始形状上下文计算出来的直方图及距离 (e) 图(a)中两个形状用文献[16]中改进的形状上下文计算出来的直方图及距离 (f) 图(a)中两个形状用本文方法改进的形状上下文计算出来的直方图及距离

2.2 背景模型和前景检测

每个像素点的背景模型由一组形状上下文直方图所组成, 即 $\{\bar{m}_{1,\chi}, \bar{m}_{2,\chi}, \dots, \bar{m}_{K,\chi}\}$, 其中 K 是直方图的个数 (3~7 个, 本文取 4 个)。每一个直方图都有一个对应的权值 $w_{k,\chi}$ 。形状上下文的计算是相当耗时的, 故在初始化的时候, 仅仅计算边缘点的形状上下文直方图, 非边缘点的形状上下文直方图直接初始化为 $\bar{0}$ 。在一幅图像中, 边缘点一般只占很小一部分, 这使得节省了大量的计算时间。

对一帧新的图像, 计算边缘点(x)的形状上下文直方图 \bar{h}_x , 并使用 χ^2 距离来匹配当前背景模型各个直方图。匹配的距离定义为:

$$D(\bar{h}_x, \bar{m}_{k,\chi}) = \min_{k'} D_{\chi}(\bar{h}_x, \bar{m}_{k',\chi}) \quad (4)$$

其中, ∂x 是 x 的一个邻域. 如果当前点的形状上下文直方图与至少一个模型直方图的匹配距离小于给定的阈值 T_F (本文的实验中取 0.4), 且该模型直方图对应的权值大于给定阈值 T_w

(本文的实验中取 0.35)，则称当前点的形状上下文直方图与背景模型相匹配，并称匹配距离最小的相匹配的背景直方图为最佳匹配直方图。若当前边缘点的形状上下文直方图与背景模型相匹配，就将该边缘分类为背景，否则分类为前景。

根据前景/背景分类的结果对当前背景模型进行更新. 若当前边缘点被分类为前景，则用当前点的形状上下文直方图替代权值最小的背景模型直方图，并将该模型直方图的权值设置为一个较小的初始权值. 若当前边缘点被分类为背景，则按下式更新最佳匹配直方图：

$$\overline{m}_{k,x} = \alpha \tilde{h}_x + (1 - \alpha) \overline{m}_{k,x} \tag{5}$$

其中 α 是设定的直方图更新速度因子(本文取 0.00001). 并更新所有背景模型直方图的权重：

$$w_{k,x} = \beta M_{k,x} + (1 - \beta) w_{k,x} \tag{6}$$

其中 β 设定的直方图权重更新速度因子(本文取 0.00001)；对于最佳匹配直方图 $M_{k,x} = 1$ ，其它的背景模型直方图 $M_{k,x} = 0$ 。

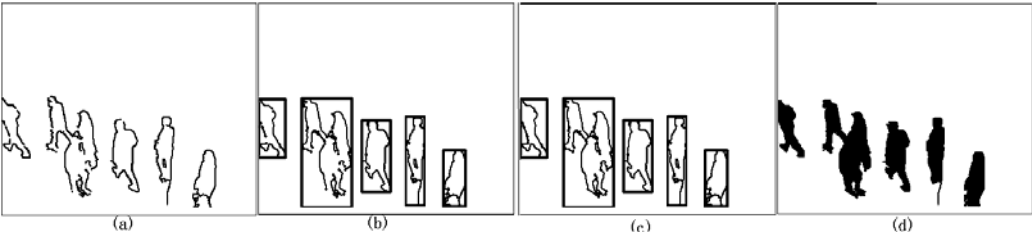


图 3 前景块生成示例 (a)检测到的前景边缘 (b)分组后的结果 (c)边缘连接后的结果 (d)填充后的结果



图 4 边缘连接操作

3 前景块的生成

以边缘形式表达的前景可以直接供基于边缘表达(edge-based)的后续处理使用. 但对于那些基于区域表达(region-based)的后续处理来说，还需要将检测到的前景转换为以区域或块(blob) 表达的形式。图 3 给出了一个生成前景块的示例. 图 3 (a) 中是以边缘形式表达的前景检测结果，由于边缘检测或前景检测时少量的漏检，有些前景物体的边缘并不封闭. 为了处理的便利，首先根据前景边缘外接矩形的临近性对前景边缘进行分组(图 3b). 根据格式塔法则(Gestalt laws)^[0] 针对前景边缘常见的 5 种断裂情形分别设计了 5 个对应的边缘连接操作(图 4). 图 3(c)是在各组内通过连接操作连接后的封闭前景边缘.填充后形成的以块形式表达前景检测结果如图 3(c)所示.

4 实验结果

本文的方法、GMM^[0]和LBP^[0]在从PETS数据集^[0] 中选取的 6 个视频序列上进行了感官和定量两种比较. 这些视频序列中包含了文献[0]中列举的大部分前景检测常见难题. 图 6 给出

了感官比较的结果. 对于 1 号而言, 由于地板的反光, GMM和LBP将大量阴影误检为了前景, 另外LBP的检测结果中很多漏检(人身上的孔洞), 而本文方法得到了跟真实数据(Ground Truth)较为接近的结果.由于缺少“干净”的背景序列进行学习(整段视频中都有前景在活动), GMM在 2 号视频上检测到了很多虚警; 同 1 号视频情况一样, LBP的检测结果中有孔洞; 本文的方法漏检了走廊尽头的小目标.虽拍摄 3 号视频的摄像机在微微抖动, 结果表明三种方法都没有收到该影响; 但由于光照的变化, GMM检测到了不少虚警; LBP检测到的结果比真实数据“臃肿”; 本文的方法得到了与真实数据几乎相同的结果. 3 号视频中光照强度快速变化, GMM基本上失效; LBP和本文方法可以抗光照快速变化. 三种方法都能应付 5 号和 6 号视频中摇摆的树枝, 但本文方法相对GMM和LBP而言, 检测到的前景形状与真实数据更为相似. 综上, 本文方法和LBP对光照变化不敏感, 而GMM则较敏感; 本文方法和LBP对“干净”训练样本的要求较GMM低; 本文方法比LBP得到的前景目标形状更为精确.为进行定量比较, 从每个视频序列中随机选取 20 帧图像进行了手工标注. 图 6 给出的定量比较结果表面本文的方法在 6 个测试视频序列上检测的平均漏检和虚警像素个数均比GMM和LBP少。

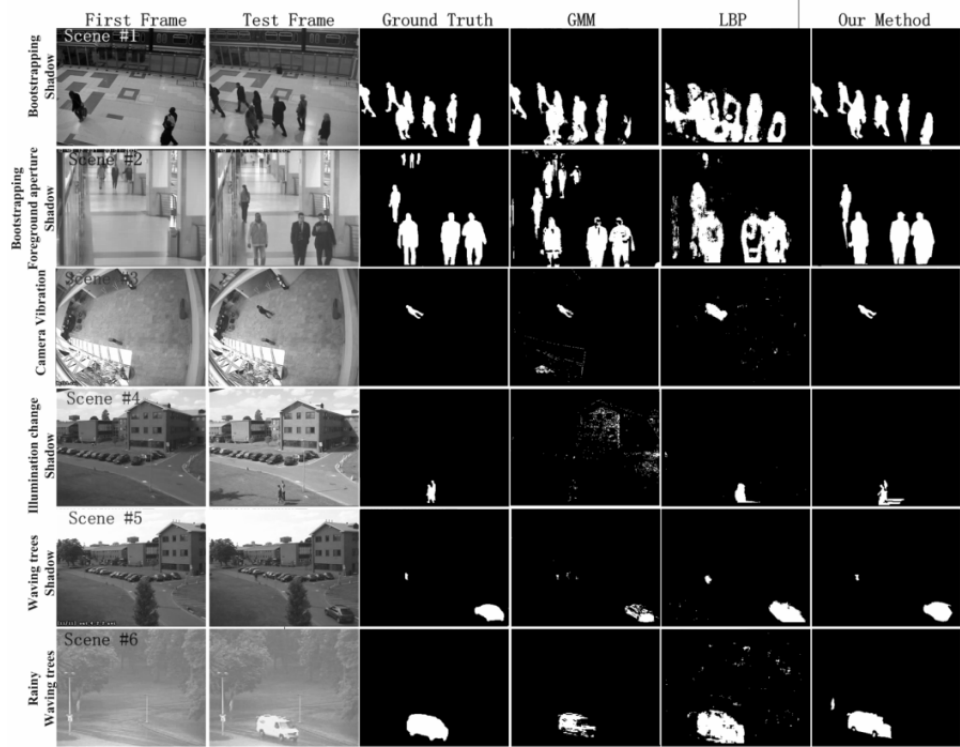


图 6 本文方法与 GMM、LBP 结果的感官比较

三种方法在 Intel 2.8GHz CUP, 2GB 内存的 PC 机上对测试视频(分辨率 360*288)的平均处理速度为: 本文方法 89.8 帧/秒, GMM 102.4 帧/秒, LBP 4.7 帧/秒. 这表明, 本文方法和 GMM 能较好满足视频监控系统的实时性要求, 而 LBP 则较难满足。

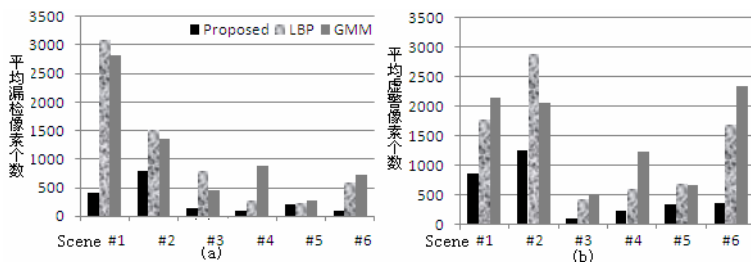


图7 本文方法与 GMM、LBP 结果的定量比较

5 结束语

本文给出了一种新颖的基于形状信息的前景检测方法. 由于形状是背景的一个非常稳定的特征. 与当前最为流行的两种前景检测方法(GMM 和 LBP)的比较实验结果表明, 本文的方法具有较好的鲁棒性, 精确性和实时性.

参 考 文 献

- [1] C. Wren and et al. Pfunder: real-time tracking of the human body [J]. PAMI, 19(7), 1997, 780–785
- [2] C. Stauffer and W.E.L.Grimson. Adaptive background mixture models for real-time tracking [A]. In: CVPR [C], vol. 2, 1999. 246–252
- [3] Z. Zivkovic. Improved adaptive Gaussian mixture model for background subtraction [A]. ICPR [C], vol. 2, 2004, 28–31
- [4] Z. Zivkovic and F. var der Heijden. Recursive unsupervised learning of finite mixture models [J]. PAMI, 26(5), 2004, 651-65
- [5] A. Elgammal, R. Duraiswami, D. Harwood, and L. Davis. Background and foreground modeling using nonparametric kernel density estimation for visual surveillance. Proc. IEEE, 90(7), 2002, 1151–1163
- [6] A. Mittal and N. Paragios. Motion-based background subtraction using adaptive kernel density estimation [A]. In: CVPR [C], 2004
- [7] O. Javed, K. Shafique, and M. Shah. A hierarchical approach to robust background subtraction using color and gradient information [A]. In: Proceedings of Workshop on Motion and Vision Computing (WMVC) [C], 2002, 22–27
- [8] S. Jabri, Z. Duric, H. Wechsler, and A. Rosenfeld. Detection and location of people in video images using adaptive fusion of color and edge information [A]. In: ICPR [C], vol.4, 2000, 627–630
- [9] M. Mason and Z. Duric. Using histograms to detect and track objects in color video [A]. In: Proc. Applied Imagery Pattern Recognition Workshop [C], 2001, 154–159
- [10] M. Heikkilä and M. Pietikäinen. A texture-based method for modeling the background and detecting moving objects. PAMI, 28(4), 2006.657–662
- [11] Y. Zhao, H. Gong and et al. Spatio-temporal patches for night background modeling by subspace learning [A]. In: ICPR [C], 2008
- [12] B. Zhang, Y. Gao, B. Zhong. Complex background modeling and motion detection based on texture pattern flow [A]. In: ICPR [C], 2008
- [13] S. Smith and J.M. Brady. SUSAN—a new approach to low level image processing [J]. IJCV, 23(1), 1997, 45–78

- [14] S. Belongie, J. Malik and J.Puzicha. Shape matching and object recognition using shape contexts [J]. PAMI, 24(24), 2002, 509—522
- [15] L.Wang, J. Shi and et al: Object Detection Combining Recognition and Segmentation [A]. In: ACCV, 2007
- [16] A. Desolneux, L. Moisan and J.M. Morel. From Gestalt Theory to Image Analysis--A Probabilistic Approach [M]. 2006, 19–37
- [17] Computational Vision Group, Reading University. PETS Dataset[EB/OL]. <http://www.cvg.rdg.ac.uk/slides/pets.html>
- [18] K. Toyama, J. Krumm, B. Brumitt, and B. Meyers. Wallflower: principles and practice of background maintenance. In: Proc. ICCV, vol. 1, 1999, 255–261

作者简介:

刘乐元 (1982—), 男, 湖北红安人, 博士研究生, 主要研究方向为模式识别、计算机视觉、智能视频监控。

桑 农 (1968—), 男, 重庆人, 博士, 教授, 博导, 主要研究方向为图像处理、模式识别、计算机视觉。

多模态导航定位中星座选择算法

邵怀宗 邵 龙 彭启琮

(电子科技大学大学通信与信息工程学院, 四川成都 610054)

摘 要: 传统 GDOP 的星座选择算法是针对单模态导航定位系统, 以各卫星伪距测量均方差相等为前提, 不适合模态间卫星伪距测量均方差不同的多模态导航定位系统。本文提出了基于加权 GDOP 的新的星座选择算法, 推导并提出了权值确定方法。仿真结果表明, 基于加权 GDOP 的星座选择算法可以明显改善定位精度。

关键词: 几何精度因子; 加权的几何精度因子; 多模态导航定位; 星座选择

A Selection Algorithm for Multi-Modal Satellite Navigation and Positioning

Shao Huai zong Shao Long Peng Qi cong

(School of Communication And Information Engineering, University of Electronic and Technology of China, ChengDu, 610054)

Abstract: Traditional constellation selection algorithms which based on the minimum geometric dilution of precision (GDOP) aim at the single modal satellite navigation and positioning system. But these algorithms are not suitable for multi-modal navigation and positioning system. This paper presents a new constellation selection algorithm which based on weighted GDOP, and the method of weight determining is deduced and presented. The simulation results indicate that the constellation selection algorithm which based on weighted GDOP can significantly improve the positioning accuracy.

Keywords: Geometric dilution of precision; Weighted GDOP; Multi-modal navigation and positioning; Selection algorithm

1 引言

提高定位精度、改善用户体验不仅是目前主流导航服务的持续要求, 也是将卫星导航定位服务推广到多样化的消费电子产品领域的重要技术保障。但目前单一依靠 GPS 系统的技术还难以满足全方位的应用需求。随着欧洲的“伽利略”系统和中国的“北斗二代”系统投入使用, 卫星导航将从 GPS 时代向 GNSS (全球卫星导航系统) 时代转变, 形成 GPS+Galileo (伽利略系统)+其它系统 (如 GLONASS、北斗, 伪卫星) 的多系统并存的局面。GNSS 在可用性、连续性和完好性方面的保障将远比单一 GPS 好。伽利略系统与 GPS 配合使用可大幅提高导航卫星的可用性, 使单一的 GPS 市区可用性从 55% 提高到 GPS/Galileo 共用时的 95%。由于伽利略系统的初始设计已充分考虑了与 GPS 的融合, 这就将形 GPS/Galileo 组合

的 GNSS 多模兼容接收机应用的批量市场，多模接收机在性价比上将大大超过单一系统的接收机。因此，多模兼容是卫星导航接收机发展的必然方向^[1]。

在导航定位应有中，假设不计卫星位置误差且各卫星具有相同的伪距测量噪声方差 σ^2 ，则用户总的定位误差为：

$$\sigma_u = \sqrt{\text{trace}(H^T H)^{-1}} \cdot \sigma = \text{GDOP} \cdot \sigma \tag{1}$$

由式(1)可以看出，用户的定位误差是伪距测量噪声均方差和 GDOP 的乘积。其中，GDOP 是由所选卫星与用户的几何分布决定的，表示了由卫星几何位置决定的对用户测距误差的放大程度，称之为几何精度因子^[2,3]。因此，合理的选择导航卫星组合使其 GDOP 最小，定位误差也就会降到最小。传统的星座选择算法主要针对单模式导航定位应用，不能直接应用与多模态导航定位系统中。为此，本文在介绍传统星座选择算法基础上，讨论了多模态导航定位应用中星座选择面临的主要问题，针对这些问题提出了多模态导航定位中星座选择算法，并仿真对比说明了该算法可以明显改善定位精度。

2 多模态导航定位系统中星座选择

2.1 传统星座选择算法

在单模式导航定位应用中，可见星数目通常为 6-12 颗。用户通常从中选择位置最好的 4 颗卫星就可以很好的完成导航定位功能。以下是常用的卫星选择方法：

(1) 从可见星中选择 4 颗进行不同的组合，分别直接计算它们的 GDOP，最小 GDOP 的组合就是最佳星座组合^[4]。

(2) 从可见星中选择 4 颗进行不同的组合，分别计算它们的组成的四面体体积，最大体积的组合就是最佳星座组合^[4,5]。

(3) 首先在用户天顶选择一颗最高的卫星，然后在最东和最北选择一颗卫星，最后从其余可见星中选择一颗组成四面体，最大体积的组合就是最佳星座组合^[6]。

(4) 从可见星中选择 4 颗进行不同的组合，采用基于雅可比算法计算求特征值的快速算法分别计算它们的 GDOP，最小 GDOP 的组合就是最佳星座组合^[7]。

在上述 4 种方法中，方法（1）可以保证选择到最佳星座组合，但是按照定义直接计算 GDOP，需要进行矩阵的转置、相乘及求逆运算，计算复杂度很大；方法（2）不需要矩阵求逆运，大大降低了计算复杂度，但是以牺牲精度为代价的；方法（3）在方法（2）的基础上，减少了组合个数，从而进一步降低了运算复杂度^[8,9]；方法（4）是在方法（1）基础上，通过计算特征值代替矩阵求逆，大大降低运算复杂度而又不牺牲定位精度，是目前在单模导航系统中比较先进的快速星座选择算法。

2.2 多模态导航定位系统中星座选择将面临的主要问题

由于单模导航定位系统一般都有 6-12 颗可见卫星，在多模态导航定位应用中，可见卫星数会相当多。根据文献[10]知道，在单模导航定位系统中选星个数大于 6 时，对定位精度改善不大^[10]。然而，随着选星数目的增加，位置解算等运算复杂度将急剧增大。

式(1)是用在单模导航定位中，假设了同一系统中伪距测量方差相同。而在多模态导航定

位应用中存在多个导航定位系统，各个导航定位系统之间的伪距测量噪声方差不同，不能直接采用式(1)来用户总的定位误差。

在多模态导航定位应用中，各个模导航定位系统的系统时钟不一致，在导航解算中需要增加解算系统间钟差未知数，因此，导航定位需要最少卫星数目会根据模态数不同而不同，每增加一个模态就需要增加一颗卫星用于解算。

综合问题（1）、（2）和（3）可知，在多模态导航定位应用中，星座选择算法是必要的，但传统的星座选择算法不能直接应用多模态导航定位中，需要提出新的星座选择算法。

2.3 加权多模星座选择算法

由多模态导航定位应用中星座选择将面临的问题（2）可知，不能直接采用式(1)来用户总的定位误差。加之，往往由于用户实际应用需要，可能并不会只根据导航定位精度，而会综合考虑如安全性、可靠性和开放性等各方面因素。例如，军事应用会优先选择本国的导航定位系统以及临时的伪卫星。基于上述原因，本文提出了采用加权的 GDOP 最小法选择卫星星座。

由于在多模态导航定位应用中存在多个导航定位系统，各个导航定位系统之间的伪距测量噪声方差不同，因此解算导航定位结果都不能直接用最小二乘法或卡尔曼滤波法，也需要用加权最小二乘的算法或加权的卡尔曼滤波算法。此时将定位方程两边同时乘以一个加权矩阵 λ ，即为：

$$\lambda Y = \lambda H X$$

其中 $\lambda = \text{diag}\left[\frac{1}{\sigma_1}, \frac{1}{\sigma_2}, \dots, \frac{1}{\sigma_n}\right]$ ， σ_i 是第 i ($i = 0, 1, \dots, n$) 颗卫星的测距均值方差值。与最小二乘法类似，加权的最小二乘算法只是将各个观测信息的观测误差方差的作为权值引入， R_n 这里表示 n 颗卫星的观测方差阵， $R_n = \text{diag}[\sigma_1^2, \sigma_2^2, \dots, \sigma_n^2]$ 。用户状态的加权最小二乘解为：

$$\hat{X}_{ls} = (H^T R_n^{-1} H)^{-1} H^T R_n^{-1} Y。其中，R_n^{-1} = \text{diag}\left[\frac{1}{\sigma_1^2}, \frac{1}{\sigma_2^2}, \dots, \frac{1}{\sigma_n^2}\right]。$$

导航系统的观测方程为： $\lambda Y = \lambda(HX + \varepsilon)$ 其中， $\varepsilon = (\sigma_1, \sigma_2, \dots, \sigma_n)^T$ 为测量噪声矢量。定位误差的最小二乘解为：

$$\begin{aligned} \Delta X &= X - \hat{X}_{ls} \\ &= (H^T R_n^{-1} H)^{-1} H^T R_n^{-1} (Y - \varepsilon) - (H^T R_n^{-1} H)^{-1} H^T R_n^{-1} Y \\ &= -(H^T R_n^{-1} H)^{-1} H^T R_n^{-1} \varepsilon \end{aligned} \tag{2}$$

定位误差的协方差为：

$$\begin{aligned} \text{cov}(\Delta X) &= E\{\Delta X \Delta X^T\} \\ &= E\left\{(H^T R_n^{-1} H)^{-1} H^T R_n^{-1} \varepsilon (H^T R_n^{-1} H)^{-1} H^T R_n^{-1} \varepsilon^T\right\} \\ &= E\left\{(H^T R_n^{-1} H)^{-1} H^T R_n^{-1} \varepsilon \varepsilon^T R_n^{-1} H (H^T R_n^{-1} H)^{-1}\right\} \\ &= (H^T R_n^{-1} H)^{-1} H^T R_n^{-1} E\{\varepsilon \varepsilon^T\} R_n^{-1} H (H^T R_n^{-1} H)^{-1} \\ &= (H^T R_n^{-1} H)^{-1} H^T R_n^{-1} R_n R_n^{-1} H (H^T R_n^{-1} H)^{-1} \\ &= (H^T R_n^{-1} H)^{-1} H^T R_n^{-1} H (H^T R_n^{-1} H)^{-1} \\ &= (H^T R_n^{-1} H)^{-1} \end{aligned} \tag{3}$$

可以得到用户总的定位误差为： $\sigma_u = \sqrt{\text{trace}(\text{cov}(\Delta X))} = \sqrt{\text{trace}((H^T R_n^{-1} H)^{-1})}$

加权 GDOP 定义如下： $\text{WGDOP} = \sqrt{\text{trace}((H^T W H)^{-1})}$ 其中， $W = R_n^{-1} = \lambda^T \lambda$ 。

多模态导航定位应用中，不同模态系统中的具有自己的权值，只需要选择合适的权值确定出加权加权矩阵 λ ，就可以计算出加权 GDOP 值，根据最小 WGDOP 只选择最佳星座组合。

3 算法仿真分析

为了分析在多模态导航定位系统中加权 GDOP 算法对用户定位精度的改善情况，本文把采用加权 GDOP 算法选星和采用传统 GDOP 算法选星的定位结果的误差进行对比分析。

选用 Navman 公司的 GPS 模块 Jupiter32 读取 GPS 原始数据导航电文。用户在地心地固坐标 (-3003243.450202, 4998042.258075, 2578507.611256) 位置某时刻观测到 8 颗可见星。现将卫星位置和对应的伪距列于表 1。

表 1 卫星位置及对应伪距数据表

卫星编号	卫星位置地心地固坐标 (x,y,z)	伪距 (m)
B1	(14217569.476371,21583589.209110,-6095682.550659)	25434136.571466
B2	(4975092.889691,25148101.388658,6967159.493564)	22112293.203265
B3	(-3329713.781625,21167998.113527,-15675958.366682)	24388843.850434
B4	(-19299993.542490,-3394821.380225,17945390.788948)	23920307.373869
G1	(-19963083.772104,7400215.518941,-15860929.636788)	25168156.471877
G2	(-3100667.629856,14942610.361375,21754151.035904)	21601466.630656
G3	(-19617516.897354,7129738.820807,16439410.594836)	21742064.377494
G4	(-25799211.944183,1221763.897234,-6174106.135327)	24709128.013688

为了分析加权 GDOP 对用户的定位精度的改善情况，做如下假设：B1,B2,B3,B4 是北斗二号卫星，G1,G2,G3,G4 是 GPS 卫星，对中国来说北斗二号的测距均值方差(1sigma)为 1.2m，GPS 的测距均值方差(1sigma)为 5m。

由于是双模，至少需要 5 颗星。采用加权 GDOP 算法选星和采用传统最高精度 GDOP 算法选星（方法（1）和（4））相关数据列于表 2。

从表 2 数据可知，当多模态中各模态之间的卫星的测距误差不同时，基于加权 GDOP 和基于传统的 GDOP 所选卫星组合不同，且基于加权 GDOP 选择卫星星座得到的总的定位误差较小。加权以后,增加了测距误差小的定位方程的比例,从而提高了定位精度；GDOP 是在等权的情况下得到的,因此定位误差较大。

表 2 基于加权 GDOP 和 GDOP 选星结果对照表

算法	卫星组合	定位误差(m)
GDOP	B1,B2,B4,G1,G3	7.541598
加权 GDOP	B1,B2,B3,B4,G3	6.028692

4 结论

传统的基于 GDOP 值的星座选择算法是在每颗卫星测距均值方差相等的条件下得到的,不适合模态之间的卫星的测距误差不同的多模态导航定位系统中。本文提出了基于加权 GDOP 值的星座选择算法,理论推导了该算法的适合多模态导航定位系统,并通过仿真对比,验证了基于加权 GDOP 选择卫星星座比基于传统 GDOP 选择卫星星座得到的总的定位误差小。同时,在单模态导航定位系统中,每颗卫星的测距均值方差相同时,基于加权 GDOP 选择卫星星座等同于基于 GDOP 选择卫星星座;每颗卫星的测距均值方差并不相同,将仿真中的卫星看成同一模态下的卫星数据,则采用本文提出的基于加权 GDOP 选择卫星星座可以进一步提高单模态导航定位系统的定位精度。因此,基于加权 GDOP 卫星星座选择算法不但可以提高多模态导航定位系统的定位精度,而且也可以改善单模态导航定位系统的定位精度。

参 考 文 献

- [1] 曹阳,朱小虎. 多模增强型卫星导航接收芯片的机遇与挑战[J].中国集成电路, 2007.1,29,64-67
- [2] Ahmed El Rabbany. Introduction to GPS The Global Positioning System[M] . London: Artech House ,2002. 155 - 159.
- [3] 张守信. GPS 卫星测量定位理论与应用[M]. 长沙: 国防科技大学出版社,1996
- [4] Milliken R J, Zoller C J. Principle of Operation of NAVSTAR and System Characteristics[J]. Navigation: Journal of The Institute of Navigation, summer, 1978, 25: 95-106
- [5] Bogen H A. Geometric Performance of the Global Positioning System. SAMSOTR-74-169, AD-783210, 1974, 1-116
- [6] Noe P S, Mayer K A, Wu T K. A Navigation Algorithm for Low-Cost GPS Receiver[J]. Navigation: Journal of The Institute of Navigation, summer, 1978, 25: 258-264
- [7] 邵龙,邵怀宗,文忠,刘波. 一种新的导航卫星星座选择算法[J]. 电子技术应用. 2008.12,34 (卷终) ,31-33
- [8] Yang Yong ,Miao Ling-juan. GDOP results in all-in-view positioning and in four optimum satellites positioning with GPS PRN codes ranging[J] . IEEE Position Location and Navigation Symposium ,2004. 723 - 727.
- [9] 郑作亚,黄城,冯初刚,张飞鹏. 4 颗卫星情况的几何优化法修正[J].天文学报,2003.8,3 (44) ,311-317
- [10] 丛丽,Ahmed I Abidat,谈展中.卫星导航几何因子的分析和仿真[J].电子学报,2006.12,12(34), 2204-2208

作者简介:

邵怀宗,男,信号与信息处理专业博士,电子科技大学副教授。其主要的研究方向是通信信号处理及其实现。在国内外重要的刊物上发表论文 50 余篇。hzshao@uestc.edu.cn

邵龙,男,2006 年于电子科技大学获得学士学位,现为电子科技大学通信与信息学院研究生。研究方向为数字信号处理。shaolong1223@gmail.com

彭启琮,男,电子科技大学大学教授、博士生导师,国家首届教学名师,主要研究方向是信号与信息处理,通信信号处理,在国内外发表重要的学术论文 120 余篇。qpeng@uestc.edu.cn

H.264/AVC的帧内模式选择优化算法¹

李仲林 胡 栋

(南京邮电大学江苏省图像处理与图像通信重点实验室, 江苏南京 210003)

摘 要: H.264/AVC 的帧内模式选择优化算法是实时编码中的一个关键技术。通过对帧内预测模式与图像原始像素方向信息相关性的研究, 结合不同预测方向上率失真的相关性, 提出了一种快速方向检测算法, 它能够代替全模式搜索而将 I4 块的 9 种预测模式减少到 4 种, 色差块的 4 种预测模式减少到 2 种。实验结果表明这种算法以很小的率失真代价实现了编码时间缩短 65% 左右。

关键词: 视频编码; 帧内模式选择优化算法; 快速方向检测; 率失真代价

Optimization Algorithm for Intra Mode Selection in H.264/AVC

Li Zhong lin Hu Dong

(jiangsu key laboratory of image processing and image communication, nanjing university of posts and telecommunications, Jiangsu Nanjing 210003, China)

Abstract: Optimization algorithm for intra mode selection in H.264/AVC is a key technology in real-time encoding. In this paper, a fast direction detection algorithm was proposed. The algorithm was based on the relevance of intra prediction mode and the direction information of the original image pixels, and the relevance of rate-distortion in different prediction directions. Instead of full intra mode search, the proposed method only searched 4 modes for 4×4 luminance blocks, and 2 modes for 8×8 chrominance blocks. The experimental results showed that the proposed algorithm can save approximately 65% of the total encoding time with little degradation in the rate-distortion performance.

Keyword: video coding; Optimization algorithm for intra mode selection; fast direction detection; R-D cost

0 引言

由联合视频组 JVT 制定的视频编码标准 H.264/AVC[1], 其帧内预测充分利用了图像的空间相关性, 根据已解码的相邻块信息来预测当前块, 并使用了率失真最优化 (RDO) 技术来进行帧内模式的选择, 使对 I 帧的编码性能达到很高。但是对每个 16×16 的图像宏块来说,

¹ 基金项目: 国家重大基础研究项目 (编号: 2007CB310607-3); 国家自然科学基金项目 (编号: GZ206005); 江苏省自然基金项目 (编号: BK2007238)

要确定出其最佳的帧内预测模式，共需要计算 592 种组合的率失真代价，且 RDO 算法的计算复杂度很高，使得编码器的编码速度很低，所以研究帧内模式选择的优化算法具有很重要的应用价值。

鉴于此，帧内模式选择的优化算法引起了广泛的研究。WANG et al. [2] 使用了 MPEG-7 中的边缘检测方法，从 4×4 的 9 种预测模式中预先选择了 4 种，大大缩短了编码时间，但是它在斜角方向上经常得出不准确的结果。TSAI et al. [3] 提出了一种基于方向检测的快速选择算法，它通过计算子块像素在不同方向上像素差的绝对值之和，来获得图像块的方向信息，然后根据方向信息来缩小待选模式的选择范围。CHENG et al. [4] 提出了一种快速三步帧内预测选择算法，它利用不同预测方向上率失真代价的相关性，排除了一些不太可能的预测模式，使待选模式从 9 种降为 6 种，极大地降低了计算复杂度。PAN et al. [5] 使用了 Sobel 算子来检测图像块的方向信息，通过它可以缩短编码时间 50% 左右，不过算法的复杂度偏高。MENG et al. [16] 通过设定的阈值来提早结束某些模式的选择，从而缩短编码的时间。

本文提出了一种帧内模式选择快速方向检测（FDD）算法，它能够以很小的率失真代价实现编码速度的大幅提升。本文的内容安排如下：第一部分简要介绍 H.264 的帧内模式选择算法，第二部分具体介绍帧内模式选择优化算法，第三部分是实验结果和分析，第四部分是结束语。

1 H.264/AVC 的帧内模式选择算法

H.264/AVC 共有三种类型的帧内预测模式：第一类是 4×4 亮度预测模式，共有 9 种预测模式；第二类是 16×16 亮度预测模式，共有 4 种预测模式；第三类是 8×8 色度预测模式，两种色度成分使用同一种预测模式，共有 4 种预测模式。

在 H.264/AVC 的官方参考模型 JM 里，对于每个 16×16 的宏块，其帧内模式的选择方法如下：首先，计算宏块 16×16 亮度 4 种预测模式的 SATD 值，选择具有最小 SATD 值的模式作为 16×16 亮度最佳预测模式，并计算选择最佳预测模式时的率失真代价；其次，将宏块分成 16 个 4×4 的子块，分别计算每个子块亮度 9 种预测模式的率失真代价，选择具有最小代价的模式作为当前 4×4 亮度最佳预测模式，并将 16 个子块的率失真代价之和作为宏块 4×4 亮度预测模式的代价；然后比较 16×16 和 4×4 亮度预测模式的率失真代价，选择具有较小代价的预测模式作为当前宏块的亮度最佳预测模式；最后，计算当前宏块对应的 8×8 色度 4 种预测模式的率失真代价，选择具有最小代价的模式作为当前宏块的色度最佳预测模式。

2 帧内模式选择优化算法

2.1 4×4 亮度块的快速方向检测算法

Tsai et al. [3] 提出了一种基于方向检测（DD）的模式选择优化算法，它通过检测出的 4×4 块的方向信息，从 9 种预测模式中选出 4~6 种作为待选模式，大大提高了编码器的编码速度。本文在这个算法的基础上，利用帧内模式在相邻预测方向上率失真代价的相关性，提出了一种快速方向检测（FDD）算法，它将待选模式固定为 4 种，进一步提高了编码速度。

FDD 算法的具体步骤如下:

Step1. 对每个 4×4 的块, 将其划分成 5 个子块, 分别求出每个子块 4 个像素的平均值, 即 $P_0 = (a+b+e+f) / 4$, $P_1 = (c+d+g+h) / 4$, $P_2 = (i+j+m+n) / 4$, $P_3 = (k+l+o+p) / 4$, $P_4 = (f+g+j+k) / 4$, 如图 1 所示。

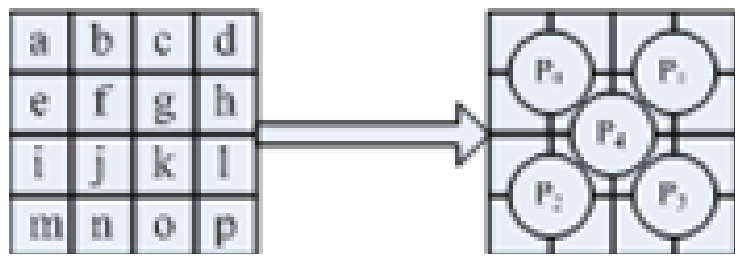


图 1 将 4×4 的块划分成 5 个子块, 并分别求其平均值

Step2. 计算 0° , 90° , 45° , 135° 四个方向上像素差的绝对值之和,
 $d^0 = |P_1 - P_0| + |P_3 - P_2|$,
 $d^{90} = |P_2 - P_0| + |P_3 - P_1|$, $d^{45} = |P_4 - P_1| + |P_2 - P_4|$, $d^{135} = |P_4 - P_0| + |P_3 - P_4|$ 。

Step3. 先从 0° 和 90° 两个方向的数值中选出具有较小数值的方向, 再从 45° 和 135° 两个方向的数值中选出具有较小数值的方向, 选择这两个方向对应的模式和夹在它们中间的方向对应的模式以及 DC 模式共 4 种作为预选模式。 4×4 块的 9 种预测模式与八个方向的对对应关系如图 2 所示。

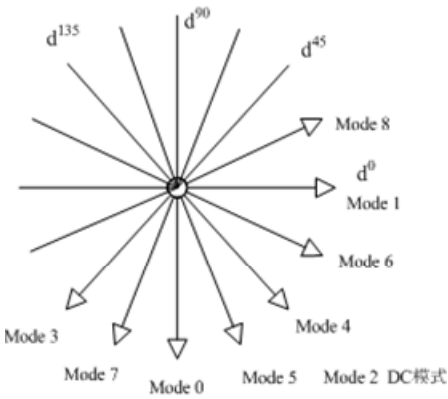


图 2 4×4 块的 9 种预测模式与方向的对应关系

Step4. 计算 4 种预选模式的率失真代价, 选择具有最小代价的模式为 14 的最佳预测模式。

2.2 8×8 色度块的快速方向检测算法

色差块的 4 种预测模式中, 有 3 种方向模式, 可通过快速方向检测算法预选出 1 种方向模式和 DC 模式一起组成待选模式, 可以很大程度的提高编码器的编码速度。

算法的具体实现步骤如下:

Step1. 对每个 4×4 的块, 将其划分成 5 个子块, 分别求出每个子块 4 个像素的平均值

$P_0 \sim P_4$ 。

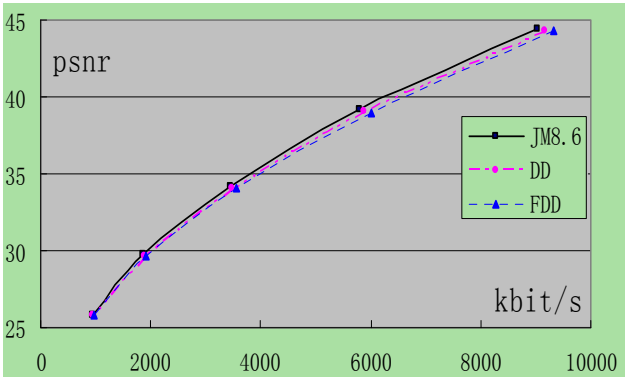
Step2. 计算 0° , 90° , 45° 三个方向上像素差的绝对值之和, $d^0 = |P_1 - P_0| + |P_3 - P_2|$, $d^{90} = |P_2 - P_0| + |P_3 - P_1|$, $d^{45} = |P_4 - P_1| + |P_2 - P_4|$ 。

Step3. 从三个方向的数值里面选出具有最小数值的方向, 然后根据所选的方向, 确定此方向对应的模式和 DC 模式作为当前块的预选模式。其中 0° 方向对应的模式为 Mode 1, 45° 方向对应的模式为 Mode 3, 90° 方向对应的模式为 Mode 2。

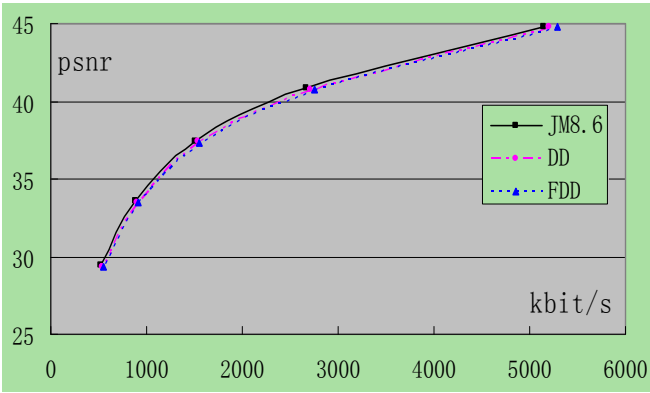
Step4. 计算 2 种预选模式的率失真代价, 选择具有较小代价的模式为色差块的最佳预测模式。

3 实验结果与分析

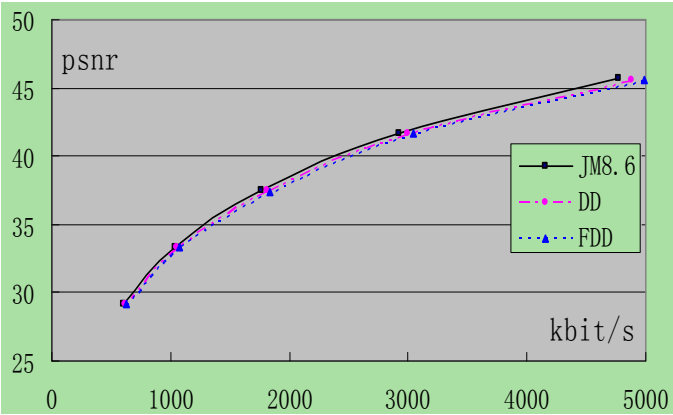
在 H.264/AVC 参考软件 JM8.6 中[7]实现了本文算法, 并在全 I 帧编码情况下对 3 个 CIF 格式的原始视频序列 bus, hall, news 的前 100 帧进行了实验, QP 分别取 18, 24, 30, 36, 42, 帧率为 30 帧/s。做仿真实验的计算机 CPU 为 AMV4000+, 内存为 1G。在图 3 中给出了三个视频序列的率失真曲线, 并在表 1 种给出了 QP=24, QP=36 时, 三种算法的仿真耗时。



(a) bus 序列的率失真曲线



(b) hall 序列的率失真曲线



(c) news 序列的率失真曲线

图 3

表 1 三种算法的仿真耗时

	QP	JM8.6	DD	FDD
bus 序列	24	1052.3s	428.8s	367.0s
	36	524.2s	209.7s	179.9s
hall 序列	24	623.9s	258.6s	219.6s
	36	373.7s	153.5s	127.9s
news 序列	24	643.3s	268.9s	227.1s
	36	384.3s	161.0s	134.0s

通过对图 3 中的 3 个序列的率失真曲线进行分析，可以看出 FDD 算法和 DD 算法的率失真曲线都十分接近于 JM8.6 的率失真曲线，可见两种优化算法的率失真性能损失都很小，FDD 算法略逊于 DD 算法的率失真性能。而对表 1 进行分析，可以明显看出 FDD 算法较 DD 算法的编码速度得到了进一步提升。

与原 JM8.6 算法相比，使用 FDD 算法后，bus, hall, news 序列的平均编码时间都减少了 65%左右，编码速度几乎提高了两倍。bus 序列的平均亮度信噪比损失为 0.11dB，比 hall, news 序列 0.05dB 的平均亮度信噪比损失略高，其原因因为在相同的 QP 时，bus 序列比其他两个序列的 I4 块要多些，而 I4 块的 FDD 算法其率失真性能会有所损失。bus, hall, news 序列的比特率增加在 3%~5%。

与 DD 算法相比，使用 FDD 算法后，bus, hall, news 序列的平均编码时间都减少了 15%左右，编码速度得到进一步提高。bus 序列的平均亮度信噪比损失为 0.03dB，比 hall, news 序列 0.01dB 的平均亮度信噪比损失略高。bus, hall, news 序列的比特率增加在 2%左右。

4 结束语

本文提出的 FDD 算法，在率失真性能降低很小的情况下，显著地提高了 H.264/AVC 的编码速度，且算法的实现简单。可是 FDD 算法在确定最佳预测模式时，还存在一些误判的情况，这种情况随着视频图像的纹理特征的变化而不同。尽量较少误判的情况，从而进一步改善算

法的率失真性能，还需要做进一步的研究。

参 考 文 献

- [1] WIEGAND T, SULLIVAN G J, BJONTEGAARD G, "Overview of the H.264/AVC Video Coding Standard", IEEE Transaction on Circuits and Systems Video Technology, vol. 13, no. 7, July 2003. [J]
- [2] WANG J. F, WANG J. C, CHEN J. T, "A Novel Fast Algorithm for Intra Mode Decision in H.264/AVC Encoders," in Proc. IEEE Int. Symp. Circuit Syst, May 2006. [J]
- [3] TSAI A. C, WANG J. F, LIN W. G, "A Simple and Robust Direction Detection Algorithm For Fast H.264 Intra Prediction", in proc. IEEE ICME, July 2007. [J]
- [4] CHENG C.C, CHANG T. S, "Fast Three Step Intra Prediction Algorithm for 4×4 block in H.264", in proc. IEEE ISCAS, Vol. 2, May 2005 [J]
- [5] PAN F, LIN X, RAHARDJA S, LIM K. P, "Fast mode decision algorithm for intra prediction in H.264/AVC video coding," IEEE Transaction on Circuits and Systems Video Technology, vol. 15, no. 7, Jul. 2005. [J]
- [6] MENG B, AU O. C, "Fast intra-prediction mode selection for 4×4 blocks in H.264," in Proc. IEEE ICASSP, vol. 3, Mar. 2003. [J]
- [7] Joint Video Team Reference Software JM 8.6. [Online] Available: <http://iphome.hhi.de/suehring/tml/download//>

作者简介:

李仲林 (1985-), 男, 山西运城人, 硕士研究生, 主要研究方向: 图像处理与多媒体通信

胡栋 (1962-), 男, 江苏南京人, 教授, 硕士生导师, 主要研究方向: 图像处理与多媒体通信

基于CORDIC算法的数字下变频器实现

王小凡 唐友喜

(电子科技大学 通信抗干扰技术国家级重点实验室, 四川 成都, 611731)

摘要: 数字下变频技术是软件无线电的关键技术之一, 其功能是将宽带高速数据流信号转变为窄带低速数据流信号, 以实现信号的实时处理。采用 CORDIC 算法实现的数字下变频器, 克服传统下变频器查询表大的缺点, 能够有效提高信号处理的效率, 减少硬件设计的代价, 非常适合在 FPGA 中实现。

关键词: 软件无线电; 数字下变频; CORDIC 算法; FPGA

Implementation of Digital Down Conversion Based on CORDIC Algorithm

Wang Xiao fan Tang You xi

(National Key Laboratory of Communication, UESTC, Chengdu 611731, China)

Abstract: Digital down conversion is one of the key technologies in software defined radio. Its function is to transfer wideband high speed data signals to narrowband low speed data signals, so that they can be processed in real time. The DDC based on CORDIC algorithm overcomes the shortcomings of traditional DDC in which a large look-up table is needed. The CORDIC-based DDC can improve the efficiency of signal processing and reduce the cost of hardware. Therefore it is very suitable for implementation in FPGA.

Keywords: software defined radio; digital down conversion; CORDIC algorithm; FPGA

1 引言

数字下变频(Digital Down Conversion, DDC)技术是软件无线电的核心技术之一^[1], 其目的是把信号搬移到更低的频率上, 将高速宽带数据流信号转变为窄带低速数据流信号, 以便实时信号处理。传统的数字下变频器由本地 NCO(Numerically Controlled Oscillator, 数控振荡器)产生的信号与输入信号相乘实现下变频, 如图 1 所示。本地 NCO 通常采用 DDS(Direct Digital Synthesize, 直接数字合成)技术, 这种方法具有结构简单, 易于实现的优点, 但若若要获得高分辨率的输出信号, 就需要大量的 ROM 资源。

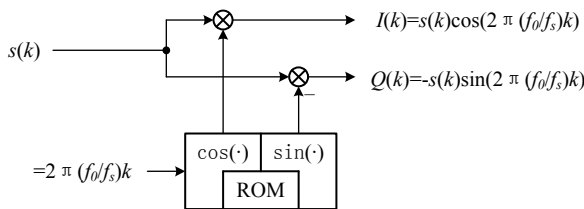


图1 数字下变频直接实现技术

本文研究的基于CORDIC(Coordinate Rotation Digital Computer, 坐标旋转数字计算机)算法实现的数字下变频技术, 不仅节约了大量的ROM资源, 而且用于混频的乘法器也可以省掉。由于CORDIC算法仅采用移位、加法和减法运算, 因此非常适合在FPGA中实现。

2 CORDIC原理

CORDIC算法是Volder^[2]在1959年提出的, 用于极坐标和笛卡尔坐标之间的转换。这是一种迭代算法, 仅需要移位、加法和减法运算。在圆周旋转模式下, CORDIC算法可以计算出一个向量旋转任意角度后的笛卡尔坐标。

为把向量 $\mathbf{v}_0=[x_0 \ y_0]^T$ 旋转一个角度 $\phi=z_0$, CORDIC算法要进行一系列逐次递减的基本旋转, 基本旋转角度为 $\phi_i=\pm\arctan(2^{-i})$, $i=0,1,\dots,n-1$ 。迭代过程中产生的中间向量的笛卡尔坐标可以用下面的等式描述

$$x_{i+1} = x_i - d_i y_i 2^{-i} \quad (1)$$

$$y_{i+1} = y_i + d_i x_i 2^{-i} \quad (2)$$

$$z_{i+1} = z_i - d_i \arctan(2^{-i}) \quad (3)$$

其中

$$d_i = \begin{cases} -1 & z_i < 0 \\ 1 & z_i \geq 0 \end{cases} \quad (4)$$

定义了每次基本旋转的方向。在经过 n 次迭代后, CORDIC算法的计算结果为

$$x_n \approx A_n [x_0 \cos(z_0) - y_0 \sin(z_0)] \quad (5)$$

$$y_n \approx A_n [y_0 \cos(z_0) + x_0 \sin(z_0)] \quad (6)$$

$$z_0 \approx 0 \quad (7)$$

其中

$$A_n = \prod_{i=0}^{n-1} \sqrt{1+2^{-2i}} \quad (8)$$

为CORDIC算法的放大因子, 它仅取决于迭代次数。CORDIC算法的计算结果是向量旋转后的放大形式。

由 $\phi = \sum_{i=0}^{n-1} \arctan(2^{-i})$ 可知, 当迭代次数 n 趋近于无穷大时, ϕ 所能覆盖的范围只有 $-99.9^\circ \sim +99.9^\circ$ 。为了使旋转角度能够覆盖 $-\pi \sim +\pi$ 整个圆周, 需要增加一次初始化旋转, 将向量先旋转 $\pm\pi/2$ 。初始化旋转可以用下面的等式描述

$$x' = -dy \quad (9)$$

$$y' = dx \tag{10}$$

$$z' = z - d \frac{\pi}{2} \tag{11}$$

其中

$$d = \begin{cases} -1 & z < 0 \\ 1 & z \geq 0 \end{cases} \tag{12}$$

经过初始化旋转后，剩余的旋转角度就被调整到 $-\pi/2 \sim +\pi/2$ 之间。

3 CORDIC算法实现数字下变频

从式(5)和式(6)可以看出，如果在每个时钟周期，用数字中频信号 $s(k)$ 代替 y_0 ，用离散相位 $\phi(k)=2\pi(f_0/f_s)k$ 代替 z_0 ，其中 f_s 为采样频率， f_0 为搬移频率，同时令 $x_0=0$ ，则CORDIC算法的计算结果相当于输入中频数据 $s(k)$ 与正弦和余弦样本值的乘积， x_n 和 y_n 分别为下变频后的正交分量 $Q(k)$ 和同向分量 $I(k)$ 的放大形式，如图2所示。 A_n 为放大因子。

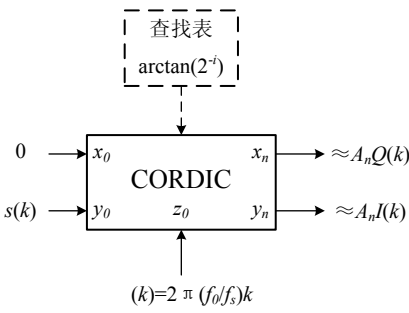


图2 基于CORDIC算法的数字下变频器

CORDIC 算法是迭代算法，如果每个时钟周期只执行一次移位和加减法运算，那么每个采样值的计算都需要 $(n+1)$ 个时钟周期，这就要求算法单元的内部时钟频率必须达到采样时钟频率的 $(n+1)$ 倍^[3],不适用于对时间要求较高的场合。为了实现高速应用的目标，可以采用流水线的结构。在这种结构中，输出数据相对于输入数据有 $(n+1)$ 个时钟周期的滞后，但在正常工作时，每个时钟周期都能输出一个数据，能够实现高速数据的实时处理，非常适合应用在数字下变频中。

在流水线结构中，每一级流水线单元主要由三个加法器组成，每一级旋转的固定基本角度 $\arctan(2^{-i})$ 通过硬线连接，省去了传统的 ROM 表。级与级之间通过寄存器相连。由于每一级流水线单元都是用简单的移位、加法和减法运算实现的，因此在整个算法过程中不需要使用任何乘法器。

4 仿真分析

采用流水线结构的CORDIC算法，所能达到的计算精度，与流水线的级数和数据的路径宽度有关。每增加一级流水线，大约可以增加一位计算精度，流水线的级数越多，所能达到

的精度越高。在定点运算中，有限字长造成的截断也会带来近似误差，数据的路径宽度越宽，截断带来的误差越小。但应该注意到，流水线的级数不应该多于数据路径的位宽，否则在超出的流水线单元中，数据经过移位和截断后将不会对计算结果产生影响。Hu^[3]对 CORDIC 算法的量化效应进行了研究，提出了达到不同精度所需的迭代次数和路径宽度组合。

对于输出精度 12 位的数字下变频器，在 Hu 研究结果的基础上，选取 14 级流水线和 18 位路径宽度。在流水线级数确定之后，放大因子 A_n 也就确定为相应的常数，因此在这里不考虑 A_n 引起的幅度变化。

设输入中频信号为频率 8MHz 的单一正弦波，采样频率 50MHz，搬移频率 1MHz，在 Modelsim 中对 Verilog HDL 建模的基于流水线 CORDIC 算法的数字下变频器进行仿真，得到下变频后 I、Q 两路信号的瞬时时域波形如图 3 所示。

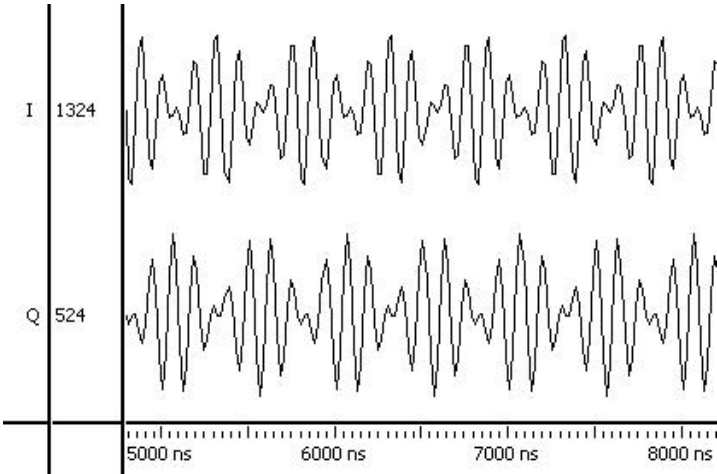


图 3 基于 CORDIC 算法的数字下变频器 ModelSim 仿真波形

为了能够对仿真结果进行直观的观察和分析，将两路仿真数据以复数的形式导入 Matlab 中，绘制出相应的功率谱密度图如图 4 所示。从图 4 可以看出，信号的频率在 7MHz，最大谐波信号幅值在-80dB 以下，完全能够满足实际应用的要求。

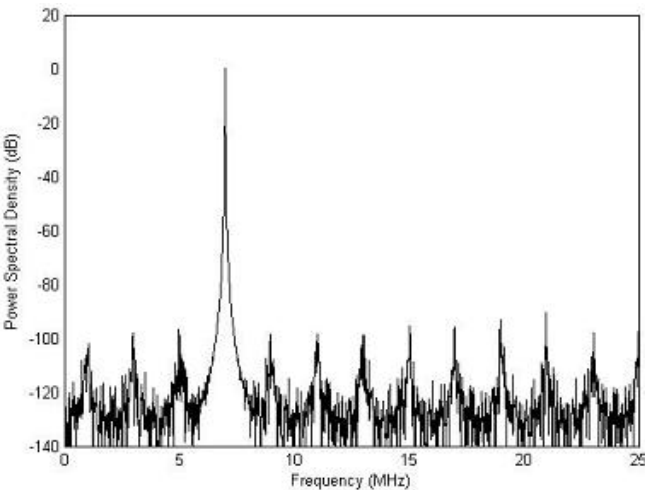


图 4 基于 CORDIC 算法的数字下变频器输出信号功率谱密度

采用 CORDIC 算法实现的数字下变频器, 输出信号在某些频点的杂散较大, 这是因为角度向量经过 $(n+1)$ 步旋转后没有与相位累加器产生的角度完全重合, 它们之间存在着相位差 δ 。对于相同的 δ , 在不同相位时产生的正弦值误差也不一样, 相位值在 0 和 π 附近最小, $\pi/2$ 和 $3\pi/2$ 附近最大。通过增加流水线级数和数据路径宽度可以减小杂散, 但同时也会引起占用硬件资源的增加, 因此需要根据实际应用的性能要求选择合适的流水线级数和数据路径宽度。

从综合报告来看, 基于流水线 CORDIC 算法的数字下变频器, 所需的硬件总量大约相当于三个相应字长的乘法器, 耗费的逻辑资源是传统下变频技术的 1.5 倍, 但由于不再需要庞大的 ROM 表, 因此节约了相当可观的硬件资源。

5 结束语

基于 CORDIC 算法的数字下变频技术, 克服了传统数字下变频器需要耗用大量 ROM 表的缺点, 减小了数字下变频器使用硬件设计时所占的资源, 提高资源的利用率, 并且能够很好的兼顾速度和精度。该数字下变频器在 Altera Cyclone II 2C35 硬件平台上工作良好。

参 考 文 献

- [1] 杨小牛, 楼才义, 徐建良. 软件无线电原理与应用[M]. 北京: 电子工业出版社, 2001. 122.
- [2] Volder J E. The CORDIC Trigonometric Computing Technique[J]. IRE Transactions on Electronic Computers, 1959, 8: 330-334.
- [3] Lohning M, Hentschel T, Fettweis G. Digital Down Conversion in Software Radio Terminals[C]. Proceedings of the 10th European Signal Processing Conference, 2000, 3: 1517-1520.
- [4] Hu Y H. The Quantization Effects of the CORDIC Algorithm[J]. IEEE Transaction on Signal Processing, 1992, 40(4): 834-844.

作者简介:

王小凡 (1984—), 男, 硕士研究生, 研究方向: 无线通信中的信号处理。

唐友喜 (1964—), 男, 教授, 博士生导师, 研究方向: 无线通信。

对角空时码在无线中继网络中的应用

赵知劲 杨进丽 王海泉

(杭州电子科技大学 通信工程学院, 浙江 杭州 310018)

摘要: 在发送端、接收端及每个中继都配置单根天线的无线中继网络中, 当发射功率极大时, 分布式空时码可以提供最大的分集度。目前用于无线网络中的空时码都被限定为酉空时码或者正交(拟正交)空时码, 那么如何将一般的空时码应用于无线中继网络中是一个很有意义的问题。现存的、高效的空时分组码一般都建立在对角空时分组码的基础之上, 而将对角空时码应用于无线中继网络中就是本文的任务。不仅如此, 本文还在极大似然解码下, 对基于成对错误概率的 chernoff 上界进行了计算并对系统进行了性能分析, 在此基础上, 提出了码的两个设计准则。

关键词: 无线中继网络; 分布式空时码; 对角空时码; 成对错误概率

Diagonal Space-Time Codes Deployed in Wireless Relay Network

Zhao Zhi jin Yang Jin li Wang Hai quan

(School of Communications Engineering, Hangzhou Dianzi University, Hangzhou 310018, China)

Abstract: It has been proved that distributed space-time codes can achieve the maximal diversity order in a wireless relay network which has a single antenna in the transmitter, in the receiver, in the relays, respectively, when the transmit power is sufficiently large. In existing works, space-time codes used in a wireless relay network are limited on unitary space-time codes or orthogonal (quasi-orthogonal) space-time codes. How to apply a general space-time code into the wireless relay network is an interesting topic. As the first step on this direction, in this paper, we apply a diagonal space-time code into the wireless relay network. Moreover, the chernoff upper bound of the pair-wise error probability (PEP) based on the maximum-likelihood decoding is calculated. Based on this bound, two design criteria are provided.

Keywords: wireless relay network; distributed space-time code; diagonal space-time codes; pair-wise error probability (PEP)

1 引言

考虑一个无线网络通信系统, 它具有 1 个发送端, 1 个接收端和 R 个中继。对于这个系统, 假定在发送端与接收端之间没有直接的通信信道, 只有借助于中继来完成发送端到接收端的通信, 并且中继之间不进行通信。另外假定发送端, 接收端及每个中继都只具有单根天线。根据在中继端进行的不同操作, 有两种协作方案: 放大——前传方案 (amplify-and-forward,

AF) 和解码——前传方案 (decode-and-forward, DF)。在 AF 方案中, 中继不需要知道发送端到接收端之间的任何信道信息也不需要知道码书信息, 它比 DF 方案简单。本文考虑前者。

对于 AF 方案, 根据在接收端所知道的信道信息不同, 提出了不同的设计方案。在文[1]中, 假定接收端知道完全的信道信息, 包括从发送端到中继的信道信息及从中继到接收端的信道信息, 作者指出如果信道的相干时间足够长, 系统的分集度可以达到中继个数 R , 并且, 基于成对错误概率证明了, 针对于多输入多输出 (Multiple Input Multiple Output, MIMO) 系统而设计的酉空时码可以提供 R 这么大的分集度, 并把这种码称为分布式空时码^[2]。此外, 在文[3]中, 正交和拟正交空时码也可以用来提供这么大分集度。

注意如上的所有应用中, 空时码都被限定为酉空时码或者正交 (拟正交) 空时码。自然地问题是, 如何将一般的空时码运用于无线中继网络中, 如黄金码 (Golden Codes)^[4]或者通过代数扩展生成的更加一般的码。本文就是针对这个问题进行研究。

作为研究如上问题的第一步, 本文将对角空时码应用于无线中继网络中。之所以选择对角空时码作为第一步是因为现存的、高效的空时分组码一般都建立在一个对角空时分组码的基础之上, 如黄金码 (Golden Codes), 理想码 (Perfect Codes)^[5], 代数空时码^[6]等。本文的主要结果就是, 把对角空时分组码应用于上述网络系统, 并给出了基于成对错误概率 (Pair-wise Error Probability, PEP) 的 Chernoff 上界的计算及系统的性能分析。并在此基础之上, 给出了码的设计准则。

本文的结构如下: 在接下来的第二节将给出文章中涉及的无线中继网络的详细描述; 在第三节中, 对角空时码被应用于无线中继网络中, 并对其性能进行分析, 在此基础之上, 码的设计准则被提出; 最后一节给出结论。

2 无线中继网络系统模型

考虑一个无线中继网络如图 1 所示。这个网络包含 $R+2$ 个节点, 其中一个发送节点 S , 一个接收节点 D , R 个中继节点, 分别为 G_1, \dots, G_R 。假设在每个节点都只有一根天线, 可以用来发送和接收信号, 但不能同时进行。并且进一步假设在发送节点 S 和接收节点 D 之间没有直接的信道, 所有的通信都要借助于中继节点。记从发送端到第 i 个中继的信道参数为 f_i , 从第 i 个中继到接收端的信道参数记为 g_i 。对上述系统做如下假设: (1) 信道为瑞利平坦衰

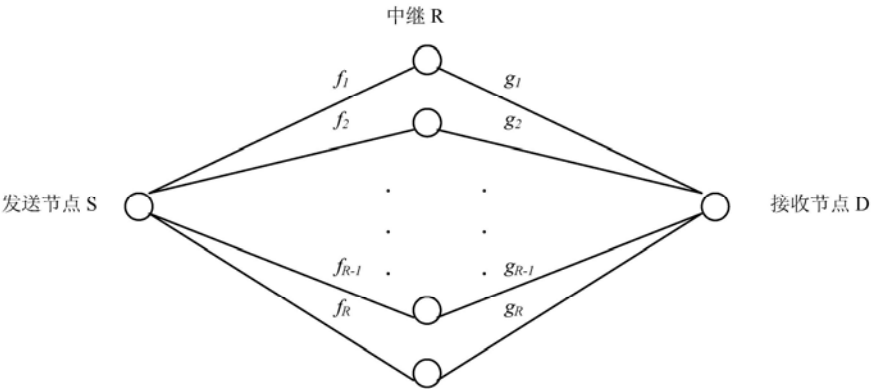


图 1 无线中继网络

落信道^[7]，衰落系数 f_i, g_i 是独立复高斯随机变量，其均值为0，方差为1；（2）所有的节点在在符号水平上是同步的；（3）相干时间间隔为 T ，即在 T 个发送过程内 f_i, g_i 为常数，到下一 T 个发送过程， f_i, g_i 变为另外独立常数。

从发送端到接收端，每一个发送过程都由两步组成。第一步，发送端发送一个长度为 T 的信息符号矢量，记为 $\mathbf{s} = [s_1, \dots, s_T]^T$ ，（假设 \mathbf{B} 为一个矩阵， \mathbf{B}^T 表示矩阵 \mathbf{B} 的转置）对 \mathbf{s} 进行能量标准化，使其满足 $\mathbf{E} \mathbf{s}^* \mathbf{s} = 1$ 。假设在发送端每个发送过程的平均功率为 P_1 ，因此，在第 i 个中继接收到的信号矢量为：

$$\mathbf{r}_i = \sqrt{P_1 T} f_i \mathbf{s} + \mathbf{v}_i \quad (1)$$

其中， \mathbf{v}_i 是在第 i 个中继端产生的标准加性高斯白噪声，即 \mathbf{v}_i 的每个元素服从均值为0，方差为1的复高斯分布 $\text{CN}(0, 1)$ 。第二步，第 i 个中继发送

$$\mathbf{t}_i = \sqrt{\frac{P_2}{P_1 + 1}} \mathbf{A}_i \mathbf{r}_i \quad (2)$$

到接收端，其中 P_2 是每个中继在每个发送过程所用的平均功率， \mathbf{A}_i 是 $T \times T$ 的矩阵，为各中继的线性加权系数矩阵。这样，在接收端，接收信号可表示为：

$$\mathbf{x} = \sum_{i=1}^R g_i \mathbf{t}_i + \mathbf{w} \quad (3)$$

其中， \mathbf{w} 是在接收端产生的标准加性高斯白噪声， \mathbf{w} 的每个元素服从均值为0，方差为1的复高斯分布 $\text{CN}(0, 1)$ 。从上面三个公式可以推知，接收端接收到的信号可以表示为如下矩阵形式：

$$\mathbf{x} = \sqrt{\frac{P_1 P_2 T}{P_1 + 1}} \mathbf{S} \mathbf{H} + \mathbf{W} \quad (4)$$

其中， $\mathbf{S} = [\mathbf{A}_1 \mathbf{s} \dots \mathbf{A}_R \mathbf{s}]$ 是一个 $T \times R$ 的矩阵，通常被称为分布式空时码； $\mathbf{H} = [f_1 g_1, \dots, f_R g_R]^T$ 为信道矩阵； $\mathbf{W} = \sqrt{\frac{P_2}{P_1 + 1}} \sum_{i=1}^R g_i \mathbf{A}_i \mathbf{v}_i + \mathbf{w}$ 为等价噪声矩阵。

这样，如果接收端知道 \mathbf{H} ，系统的极大似然（maximum-likelihood, ML）解码就成为

$$\arg \min_s \left(\mathbf{x} - \sqrt{\frac{P_1 P_2 T}{P_1 + 1}} \mathbf{S} \mathbf{H} \right)^* [\text{Cov}(\mathbf{x})]^{-1} \left(\mathbf{x} - \sqrt{\frac{P_1 P_2 T}{P_1 + 1}} \mathbf{S} \mathbf{H} \right) \quad (5)$$

对于上述系统，一个主要的问题是如何设计矩阵 \mathbf{A}_i ，使得在ML解码下，系统的错误概率达到最小。文章[1][3]中，假设矩阵 \mathbf{A}_i 是酉矩阵或是正交矩阵的条件下，给出了基于成对错误概率（PEP）的Chernoff上界的计算及系统的性能分析，并得到了矩阵 \mathbf{A}_i 的设计准则。从这些结果来看，针对一个多天线系统而设计的空时分组码可以恰到好处地应用于上述 \mathbf{A}_i 的设计。这样，自然地问题是，在文章[1][3]中，矩阵 \mathbf{A}_i 是酉矩阵或是正交矩阵的假设是否是必要的。或者说，如果把一般的空时分组码应用于上述网络系统，结果会是如何。

现存的、高效的空时分组码一般都建立在对角空时分组码的基础之上，如黄金码（Golden Codes），理想码（Perfect Codes），代数空时码等。本文的以下部分，就试图把对角空时分组码应用于上述网络系统进行码的设计，即利用对角空时分组码来设计矩阵 \mathbf{A}_i ，着重做出系统的性能分析。有关对角空时分组码的内容，由于篇幅的限制，这里省略其介绍。

3 基于对角空时码的分布式空时码及其性能分析

假定 $T=R=n$ 。一个对角空时码可以表示为 $\mathbf{C} = \text{diag}\{x_1, \dots, x_n\}$ ，这里，

$$[x_1, \dots, x_T]^T = \mathbf{G}[s_1, \dots, s_T]^T$$

并且， \mathbf{G} 是一个 $n \times n$ 复矩阵， s_i 取自某个星座，例如，QAM 或 PSK 等。令矩阵 \mathbf{A}_i 的第 i 行等于矩阵 \mathbf{G} 的第 i 行，其余部分全设为 0。这样，就可利用一个对角空时码构造出一个分布式空时码。例如， $T=R=n=2$ ，对应于黄金码的对角空时分组码为：

$$\mathbf{C} = \frac{1}{\sqrt{5}} \begin{bmatrix} \alpha & \alpha\theta \\ \bar{\alpha} & \bar{\alpha}\bar{\theta} \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \end{bmatrix},$$

其中， $\theta = \frac{1+\sqrt{5}}{2}$ ， $\bar{\theta} = \frac{1-\sqrt{5}}{2}$ ， $\alpha = 1+i-i\theta$ ， $\bar{\alpha} = 1+i-i\bar{\theta}$ 。很明显，利用上述方法，

$$\text{我们有： } \mathbf{A}_1 = \frac{1}{\sqrt{5}} \begin{bmatrix} \alpha & \alpha\theta \\ 0 & 0 \end{bmatrix}, \quad \mathbf{A}_2 = \frac{1}{\sqrt{5}} \begin{bmatrix} 0 & 0 \\ \bar{\alpha} & \bar{\alpha}\bar{\theta} \end{bmatrix}.$$

下面我们就上述得到的分布式空时码作性能分析。在这由于得到错误概率公式极其困难，我们来考虑系统的成对错误概率（PEP）。

假设接收端知道完全的信道信息 f_i 及 g_i 。因此，给定一组 f_i 及 g_i ，可以看出， \mathbf{W} 是一个高斯随机矢量，其均值为 $\mathbf{0}$ ，协方差为 $\frac{P_2}{P_1+1} \sum_{i=1}^R |g_i|^2 \mathbf{A}_i \mathbf{A}_i^* + \mathbf{I}_T$ 。假设在发送端发送的信号为 \mathbf{s}_k ，则由（4）式，接收端接收到的信号 \mathbf{x} 也是高斯随机矢量，均值为 $\mathbf{0}$ ，协方差为 $\frac{P_2}{P_1+1} \sum_{i=1}^R |g_i|^2 \mathbf{A}_i \mathbf{A}_i^* + \mathbf{I}_T$ 。因此，在给定 f_i ， g_i 和 \mathbf{s}_k 的条件下， \mathbf{x} 的概率密度函数为：

$$P(\mathbf{x}|\mathbf{s}_k) = \frac{e^{-\left(\mathbf{x} - \sqrt{\frac{P_1 P_2 T}{P_1+1}} \mathbf{s}_k \mathbf{H}\right)^* [\text{Cov}(\mathbf{x})]^{-1} \left(\mathbf{x} - \sqrt{\frac{P_1 P_2 T}{P_1+1}} \mathbf{s}_k \mathbf{H}\right)}}{\pi^T |\text{Cov}(\mathbf{x})|} \quad (6)$$

其中， $\text{Cov}(\mathbf{x}) = \frac{P_2}{P_1+1} \sum_{i=1}^R |g_i|^2 \mathbf{A}_i \mathbf{A}_i^* + \mathbf{I}_T$ 。这样，容易得到系统的极大似然（ML）解码为：

$$\arg \max_{\mathbf{s}_k} P(\mathbf{x}|\mathbf{s}_k) = \arg \min_{\mathbf{s}_k} \left(\mathbf{x} - \sqrt{\frac{P_1 P_2 T}{P_1+1}} \mathbf{s}_k \mathbf{H} \right)^* [\text{Cov}(\mathbf{x})]^{-1} \left(\mathbf{x} - \sqrt{\frac{P_1 P_2 T}{P_1+1}} \mathbf{s}_k \mathbf{H} \right) \quad (7)$$

假定系统只有两个码字 \mathbf{s}_k 和 \mathbf{s}_l ，则我们有下列结论：

定理 1 对于上述系统，假定 $P_1 = P/2$ ， $P_1 = P/(2R)$ ，则发送 \mathbf{s}_k 而 ML 译码器译为 \mathbf{s}_l 的成对错误概率（PEP）的 Chernoff 上界为：

$$P_{kl} \leq \left[\frac{8n}{8n + \sigma^2 PT} + \frac{8\sigma^2 PTR}{(8n + \sigma^2 PT)^2} \log \left(P + \frac{8n}{\sigma^2 T} \right) \right]^{\text{rank } \mathbf{M}} \quad (8)$$

其中， n 为 n_1, \dots, n_R 中的最大值， n_i 为 $\mathbf{A}_i \mathbf{A}_i^* = \text{diag}\{0, \dots, 0, n_i, 0, \dots, 0\}$ 的对角线上的第 i 个值， σ^2 为 \mathbf{M} 的最小的非零奇异值，及 $\mathbf{M} = (\mathbf{S}_k - \mathbf{S}_l)^* (\mathbf{S}_k - \mathbf{S}_l)$ 。

定理的证明可参考文献[1]。

为了极小化(8)式中的上界,很显然,首先, \mathbf{M} 的秩 $\text{rank}\mathbf{M}$ 应该达到最大,其次, \mathbf{M} 的最小的非零奇异值 σ^2 也应该越大越好。这样,我们得到对角的分布式空时码的设计准则:

- (1) 对于所有的 $k \neq l$, $\mathbf{M} = (\mathbf{S}_k - \mathbf{S}_l)^* (\mathbf{S}_k - \mathbf{S}_l)$ 必须为满秩矩阵,以获得满分集增益;
- (2) 对于所有的 $k \neq l$, \mathbf{M} 的非零奇异值的最小值必须越大越好,以获得较高的编码增益;

4 结论

本文提出将对角空时码应用于无线中继网络中进行码的设计,完成了将一般的空时码运用于无线中继网络中得第一步。在接收端知道完全的信道信息的条件下,研究了系统的最大似然解码,计算了成对错误概率的 chernoff 上界并对其进行了分析,得出了对角空时码的两个设计准则:秩准则和奇异值准则。接下来的工作是根据得到的这两个码的设计准则来设计对角空时码,找到性能比较好的码字进行仿真,并与酉空时码和正交(拟正交)空时码比较。

参 考 文 献

- [1] JING Y, HASSIBI B. Distributed space-time coding in wireless relay networks [J]. IEEE Transactions on Wireless Communications, vol: 5, NO: 12, pp: 3524–3536, December, 2006.
- [2] LANEMAN J N, WORNELL G W. Distributed space-time-coded protocols for exploiting cooperative diversity in wireless network [J]. IEEE Transactions on Information Theory, vol: 49, NO: 10, pp: 2415–2425, October, 2003.
- [3] JING Y, JAFARKHANI H. Using orthogonal and quasi-orthogonal designs in wireless relay networks [J]. IEEE Transactions on Information Theory, vol: 53, NO: 11, pp: 4106–4118, November, 2007.
- [4] BELFIORE J C, REKAYA G, VITERBO E. The Golden code: a 2×2 full-rate space-time code with non-vanishing determinant [J]. IEEE Transactions on Information Theory, vol: 51, NO: 4, pp: 1432–1436, April, 2005.
- [5] OGGIER F, REKAYA G, Member, IEEE, BELFIORE J C. Perfect space-time block codes [J]. IEEE Transactions on Information Theory, VOL:52, NO: 9, September, 2006.
- [6] SUSINDER RAJAN G and SUNDAR RAJAN B. Algebraic distributed space time codes with low decoding complexity [J]. Transactions on Wireless Communications, vol: 7, NO: 10, pp: 3962-3971, October, 2008.
- [7] JAFARKHANI H. Space-Time Coding: Theory and Practice[M]. 西安, 西安交通大学出版社, 2007.

作者介绍:

赵知劲 (1959-), 女, 浙江宁波, 博士, 教授、博士生导师, 通信信号处理、自适应信号处理、认知无线电等;

杨进丽 (1986-), 女, 山东临沂, 硕士研究生, 空时码及无线网络中码的设计;

王海泉 (1964-), 男, 浙江义乌, 博士, 讲师, 多天线系统及空时码的设计与应用。

认知无线电网络中一种自适应TDD传输策略的研究

陈晓辉 孟庆民 郑宝玉

(南京邮电大学, 通信与信息工程学院, 江苏 南京, 210003)

摘要: 认知无线电(CR, Cognitive Radio)技术是人们用来改变现有频谱短缺局面的一项关键技术。802.22 协议在 MAC 核心内容是共存性机制 (Coexistence), 基于 OFDMA/TDD 技术。受 YING-CHANG LIANG 在解决农村网络中传输时延问题^[1]的启发, 本文提出了一种自适应时分双工技术, 利用下行和上行的传输时隙传输额外的上行数据, 极有效地提高了在范围较大的农村网络中的传播效率, 计算机仿真结果显示, 当使用该技术时, 上行吞吐量增益提高明显, 最多可达约 30 %

关键词: 认知无线电; 802.22; 正交频分多址接入; 自适应时分双工; 时分双工

An Adaptive TDD Transmission Strategy In Cognitive Radio Network

Chen Xiao Hui Meng Qing Min Zheng Bao Yu

Abstract:: CR (Cognitive Radio) is a new technology to effectively change the current spectrum shortage situation. The core content of the MAC mechanism of 802.22 protocol is coexistence, which based on OFDMA/TDD technology. We proposed adaptive TDD technology, which was lightened by YING-CHANG LIANG resolving the transmit delay. The technology allows the transition gap between DL and UL subframes to be CPE-dependant, effectively increases the efficiency of large coverage, rural network. The gain in average UL capacity when employing the proposed adaptive TDD scheme is significant. The highest gain is around 30 percent, and the lowest gain is around 5 percent.

Keyword: CR; 802.22; OFDMA; ATDD; TDD

1 引言¹

随着信息时代的到来, 无线设备被广泛使用, 使得无线频谱成为现代社会不可或缺的甚至是比土地和矿藏更宝贵的资源。但是如果对城市地区和乡村偏远地区进行无线电频谱扫描, 就会发现: 某些频段在大部分时间里是空闲的, 一些频段只是有时被占用, 剩余频段则被过度使用。出现这种现象的一个原因是目前有线电视被广泛使用, 收看无线电视的观众大大减少, 在城市地区这一现象尤其突出。

¹ 基金项目: 江苏省自然科学基金重点项目 (BK2007729)。

认知无线电(CR)^[2]被认为是解决上述无线频谱低利用率问题的最佳方案。认知无线电是一种智能的无线通信系统，它能感知周围无线环境，通过对环境的理解、主动学习等措施，实现在特定的无线操作参数上(如功率、载波调制和编码等)的实时改变并且能调整系统的内部状态，适应外部无线环境的变化。2004 年 10 月，IEEE 802 委员会成立了 802.22^{[3] [4] [5]}工作组，授权开发一个共同操作的点到多点(P-MP)的空中接口(即物理层和 MAC 层)标准，该标准用于现存广播电视服务所在的频段，实现基于认知无线电的无线区域网(WRAN)。802.22 协议决定了多个设备之间的协同操作。协议交互发生在协议的每一层。其中 802.22 MAC 层协议栈对于所有被支持的服务都是相同的。802.22 中 MAC 层的中心目的是共享无线信道资源。MAC 协议定义了如何和何时初始化在信道上的传输。

MAC 协议中有 3 个关键性的问题：(1)接入初始化(2)感知管理和频谱管理(3)传播延迟的解决。现有文章绝大多数基于前两个问题的讨论。本文主要是提出一种新的自适应时分双工方式(ATTDD)增大系统的吞吐量，改进传统方案，计算机仿真显示传输时延在一定程度上得到了改善。在解决农村这种传播范围较大的地区^{[6] [7] [8]}的传输时延问题很有效果。

2 传统的OFDMA/TDD方式

一个 OFDMA 帧是一个时频二维结构^[9]，图 1 所示的是一个 OFDMA/TDD 方式下的帧结构，纵轴单位为子信道，横轴单位为符号。MAC 层调度的最小单位为一个子信道和一个或多个符号组成的时隙，具体符号数由子载波映射方式决定。每帧从时域上被分为上行(UP-Link)和下行(Down-Link,DL)两部分。上行和下行的分界是发送/接收转换间隔(Transmit/Receive Transition Gap,TTG)和接收/发送转换间隔(Receive/Transmit Transition Gap,RTG)：每个下行帧的开始为前同步符号，紧跟其后的是帧控制头(Frame Control Header,FCH)、下行映射(Down-Link Mapping,DL-MAP)和上行映射(Up-Link Mapping,UL-MAP)。

帧控制头中规定了下行映射采取的调制编码方式和长度。下行映射紧跟在帧控制头之后，上行映射紧跟在下行映射之后。下行的突发排列方式和采用的编码调制方式由下行映射指定，相应的上行编码调制方式由上行映射制定。

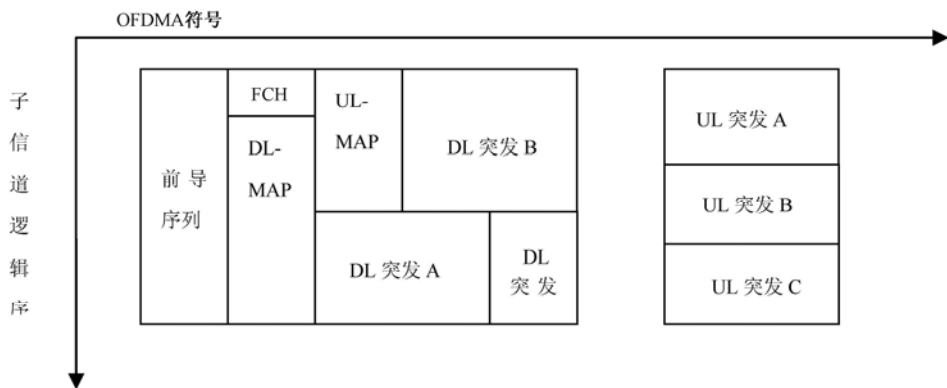


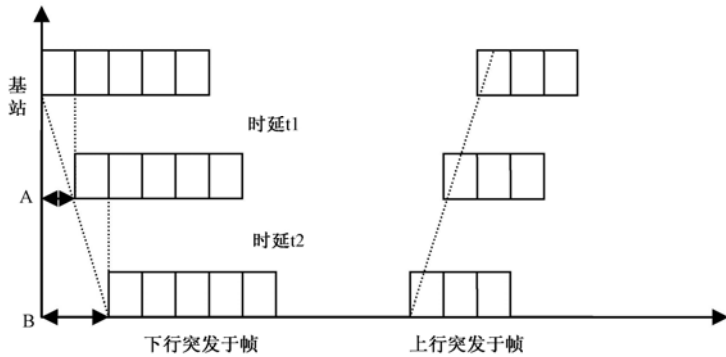
图1 OFDMA 帧格式

基站和客户端之间的距离不等，从基站到客户端设备的传播延迟自然不同，这样一来，

不同的客户端设备在不同的时隙完成下行接收。如图所示，一个近距离的客户端 A 可以比一个远的客户端 B 早很多完成它的下行接收，那么 A 自然要比 B 早很久准备开始上行传输。然而，由于传统的 OFDMA/TDD 方式对同步的要求非常严格，为了在基站保证可靠的接收，从不同的客户端设备传过来的上行突发传输必须在基站处让 OFDMA 的符号在边界上对齐。也就是说，A 在收到下行接收准备开始上行传输到它真正开始上行传输中的那断等待时间被白白浪费掉了。上述情况可用图 2 说明，图中，近距离的客户端设备 A 发出的上行传输在排列上被推迟（在基站中），目的是为了等待远距离客户端设备 B 发来的上行传输。近距离客户端设备 A 以这种等待延迟的方式消除了客户端 A 和客户端 B 的传播时间往返差距。

以上是只有 2 个客户端的情况，对于多个客户端，现在的方法以所有客户端设备中的上行传输最远的那个为标准。具体来说，即使在近距离客户端设备完成下行接收，并已经准备好进行上行传输，它还是要继续等待，以便和最远的客户端设备上行传输在到达基站时 OFDMA 符号能够对齐。

802.22 农村网络针对大范围服务时，覆盖半径将达到 100 公里以上。这时，近距离和远距离客户端在往返传播产生的差异造成的时延将会十分显著。例如，如果覆盖半径为 100 公里，那么环形传播产生的差异造成的时延将达到 660 微秒。这种时延，在典型的帧持续时间 5 至 10 毫秒的传输下将会十分明显，而且会严重降低 802.22 无线局域网的运营效率。



3 改进的自适应时分双工方式（ATTD）

我们提出的方案，称为自适应时分双工（ATTD），它利用下行与上行子帧之间的传输时隙传输额外的上行数据。具体地说，完成其下行的接收后，近距离的客户端设备首先按计划开始上行传输，接着远距离客户端设备开始上行传输。在这些进行的时候，所有客户端设备的 OFDMA 的符号边界都在基站实行同步以保证可靠的通信。

自适应时分双工可以用图 3 说明。当拖延近距离客户端设备 A 和远距离客户端设备 B 之间的往返传播产生的差异与 OFDMA 的符号时间具有可比性时，客户端设备 A 在完成下行接收后提前进行上行传输而不用等待远距离的客户端设备 B。B 定于晚些时候开始进行上行传播，它传播的时候所有的客户端设备的 OFDMA 符号边界都已在基站中被排列好了。图 3 中，在基站里，客户端设备 B 的第一个 OFDMA 符号与客户端设备 A 的第 2 个 OFDMA 符号对齐。这其实相当于在同样的时间内，近距离客户端 A 利用等待时间多传输了一个 OFDMA 符号，如图 3 中阴影部分所示。

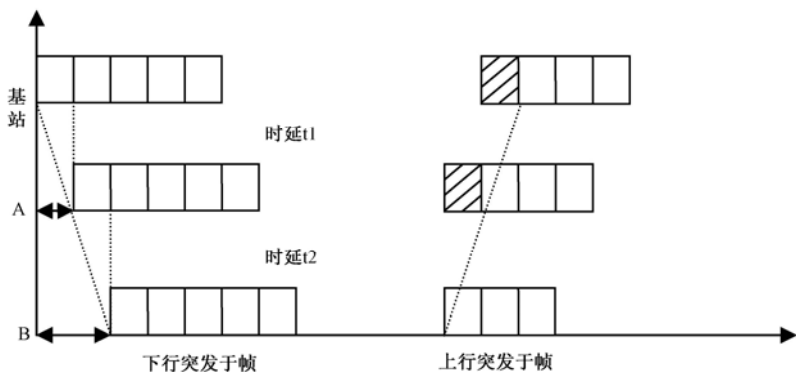


图 3 改进过的 OFDMA/ATDD 传输示意

当自适应 TDD 的允许近距离客户端设备提前启动上行传输并得到更多的 OFDMA 符号，这吞吐量增益有利于所有该系统下的用户，对上行链路及下行链路使用也是有利的。这是因为基站可以以不同的方式利用自适应 TDD 的吞吐量增益，例如，增加其中某些用户的上行带宽，或增加上行用户数量，甚至有可能在保持上行负载不变的情况下安排更多的下行链路传输。由于 802.22 帧结构的灵活性，上述情况都可以画出不同的帧结构图。下文的仿真中仅考虑增加上行传输量，至于是增加用户数量还是增加上行用户带宽可以按照实际需要自由设定。

图 4 反映了上述论点，例如常规 TDD 和自适应 TDD 的带宽分配情况显示。当使用常规 TDD 时，基站可以为 3 个用户安排下行和上行的传输。当使用自适应 TDD 时，两个近距离的用户，即用户 A 和 C 可以比其余用户早传输上行数据的 1 个符号。由于这 1 个 OFDMA 符号的增益，在分配给现有的 3 个用户的带宽不变的基础上，基站可以利用原先空等增加出来的时隙传输额外的数据，如图阴影部分所示：若某个已有用户上行突发量大时也可以用来给它分配额外的带宽；也可以用来传输一个额外的用户，若是这样只需在下行帧上加上一个 OFDMA 符号即可，如图虚线部分所示。

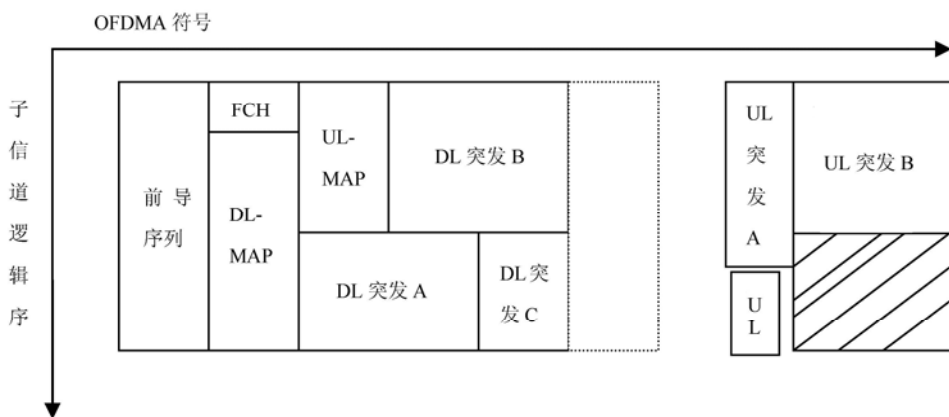


图 4 使用 ATDD 方式的帧结构

4 仿真结果数据分析（数据上行突发情况）

为了证明自适应时分双工的性能增益，我们认为这是一个 802.22 情况下的部署，考虑到

802.22 和 802.16e 参数的某些相似性, 本文参考了相关文献^{[10] [11]}中关于 OFDM 和 OFDMA 帧长度、保护时间以及子载波数的设定。在仿真环境中, 本文将区域半径设为 50 公里; 所有距离基站 5 公里内的客户端被视为近距离客户端设备, 其余不在此区域半径范围内的则被视为远距离客户端设备。我们把远距离客户端设备各按不同的比例安置, 从百分之四十直到到百分之九十, 近距离客户端设备比例的高低, 分别代表了客户端设备分布分散和集中的状况。我们假定所有客户端均表现出良好的信道条件状况, 在这里, 我们假定所有客户端设备都采用 64 QAM 调制和 3/4 码速率进行传输。我们把帧长分为 5, 10 和 20 毫秒。信道带宽是 6M Hz, OFDMA 基于 1024 快速傅立叶变换 (FFT) 的大小, 循环前缀均设为 1/8。

图 5 中, 我们绘制了远距离客户端设备相应的上行吞吐量的百分比增益。在所选参数的条件下, 一个位于基站辐射范围 5 公里内的近距离客户端设备可以比不在此范围内的远距离客户端提前一个 OFDMA 符号传输上行链路数据。采用 OFDMA/ATDD 的方法, 上行链路传输能力的平均增益相当显著。当额外 OFDMA 符号获得增益的数量上升到一定时, 上行链路传输能力的百分比增益便取决于帧的大小。图中, 实线、虚线和点化线分别代表 5ms、10ms 和 20ms 的帧取得的百分比增益。从图中可以明显看出, 帧越短, 上行链路传输的百分比增益越高; 远距离客户数越多, 上行吞吐量增益越大。当远距离客户端用户达到 90% 时, 使用 5ms 的帧上行吞吐量增益最大, 大约可以提高百分之三十; 当远距离客户端用户占 40% 时, 使用 20ms 的帧长涨幅最低, 约百分之五。远距离的用户端设备的比例增加时, 增益将逐渐增大的这种趋势可进行如下解释: 上行链路吞吐量的绝对增益几乎是常数 (由于每一个 OFDMA 符号的增益是固定的), 而另一方面, 绝对平均吞吐量随着远距离客户端设备的增加而减少。其结果是, 百分比增益, 即绝对增益除以绝对吞吐量, 将随着远距离客户端设备的增加而增加。

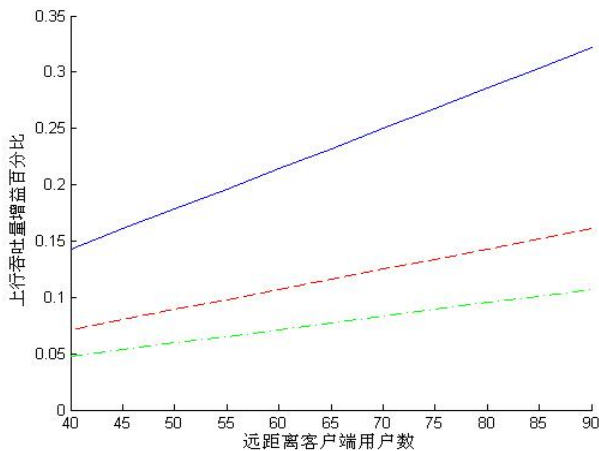


图 5 采用自适应时分双工产生的吞吐量增益

5 结束语

本文提出了一种自适应时分复用 (ATDD) 技术有效地改善了 802.22 中传输范围很大的农村网络的传播时延问题。但在实际应用中, 该模型还有待进一步的完善。如本模型并没有

考虑实际信道的不稳定性,如传输错误重传的情况,还有如 5ms 的帧里数据信息过少等问题,都有待进一步的探讨。

参 考 文 献

- [1] Cognitive radio on TV bands: a new approach to provide wireless connectivity for rural areas ,Ying-Chang Liang; Anh Tuan Hoang; Hsiao-Hwa Chen; Wireless Communications, IEEE Volume 15, Issue 3, June 2008 Page(s):16 - 22 ,Digital Object Identifier 10.1109/MWC.2008.4547518.
- [2] J. Mitola, “Cognitive Radio for Flexible Mobile Multi-media Communications,” Proc. IEEE Int’l. Wksp. Mobile Multimedia Commun., 1999, pp. 3–10.
- [3] IEEE 802.22 Wireless RAN, “Functional Requirements for the 802.22 WRAN Standard, IEEE 802.22-05/0007r46,” Oct. 2005.
- [4] Y.-C. Liang et al., “System Description and Operation Principles for IEEE 802.22 WRANs,” <http://www.ieee802.org/22/>, Nov. 2005.
- [5] IEEE 802.22 WG Web site; <http://www.ieee802.org/22/>.
- [6] Y. Kawasumi, “Deployment of WiFi for Rural Communities in Japan and ITU’s Initiative for Pilot Projects,” Proc. 6th Int’l. Wksp. Enterprise Networking and Computing in Healthcare Industry, 2004, HEALTHCOM2004, June 2004.
- [7] M. D. Farrimond, “PCN and Other Radio-Based Telecommunications Technologies for Rural Regions of the World,” Proc. 2nd Int’l. Conf. Rural Telecommun., London, U.K., 1990, pp. 99–104.
- [8] R. Westerveld and R. Prasad, “Rural Communication in India Using Fixed Cellular Radio Systems,” IEEE Commun. Mag., Oct. 1994, pp. 70–74.
- [9] WiMAX 技术应用详解, 刘波, 安娜, 黄旭林编著, 人民邮电出版社, 2007 年 8 月出版。
- [10] 802.16 宽带无线城域网媒体接入控制协议研究与系统实现 周继华 石晶林, 移动通信, 2007 年 07 期.
- [11] Rural Net 802.11-Based Low-Cost Networking for Rural India; <http://www.cse.iitk.ac.in/users/braman/dgp.html>.

作者简介:

陈晓辉(1984—), 男, 江苏淮安人, 南京邮电大学硕士研究生, 研究方向为无线通信与信号处理技术。

孟庆民(1965—), 男, 江苏滨海人, 南京邮电大学博士后, 现主要研究方向为无线 mesh 网协同策略设计、MIMO 网络和宽带多载波与多址系统性能评估、多跳中继和物理层网络编码, 无线 mesh 网和认知无线网中资源管理。

郑宝玉(1945—), 男, 福建闽侯人, 南京邮电大学教授, 博导, 上海交通大学兼职教授, 博导, 研究方向为无线通信与网络信号处理, 现代通信中的智能信号与信息处理等。

一种基于延时的AODV路由协议的改进

陈哲愚 张 建 陈 燕

(南京邮电大学 信号处理与传输研究院, 江苏 南京 210003)

摘 要: 移动 Ad hoc 网络是一种动态变化的网络, 其路由协议的目标是快速、准确、高效与可扩展。AODV 作为典型的 Ad hoc 路由协议, 能够较充分的利用无线网络资源, 具备较强的移动性和业务量的适应能力。针对 AODV 在高负载下性能快速下降的问题, 提出了一种改进协议 AODV-D, 该协议能够获取节点相邻链路的延时, 并使用基于平均延时的泛洪算法进行路径选择。NS2 仿真表明 AODV-D 在高负载下能很好的改进原协议的延时与路由负荷等性能。
关键词: 通信网络技术; AODV-D; 泛洪算法; 延时; NS-2

An Improvement of AODV Based on Delay

Chen Zhe yu Zhang Jian Chen Yan

(Institute of Signal Processing and Transmission, Nanjing University of Posts & Telecommunications, Nanjing 210003, P. R. China)

Abstract: MANET(Mobile Ad hoc Network) is a kind of dynamic network, with the objective of its routing protocol to be fast, accurate, efficient and scalable. AODV (Ad hoc On Demand), as a typical Ad hoc routing protocol, is able to make full use of the network resource, so that it owns relatively strong mobility and adaptability. This paper proposed AODV-D (AODV Based on Delay), an improved protocol, which can obtain delay of the link to neighbors and select the path through flooding algorithm based on delay, to solve performance decline problem of AODV under high network load. NS2 simulation demonstrated that the performance of delay and route load is improved by the new protocol.

Keywords: communication network technology; AODV-D; flooding algorithm; delay; NS-2

0 引言

移动 Ad Hoc 网络^[1,2,3]由一组带无线收发装置的移动节点组成,不依赖固定的基础设施,是一种典型的动态网络。组成移动 Ad Hoc 网络的节点既是发送者也是接收者,当不在彼此通信范围内时,可通过其他节点中继实现多跳通信。由于移动 Ad Hoc 网络具有分布性、抗毁性、易构性等特点,而被广泛用于战场通信、地震、抢险等紧急场合。

移动 Ad hoc 网络路由协议一直是应用研究的重点,也是难点,涉及到网络层、MAC 层、

物理层多个层次的跨层设计，QoS 保证、功率控制、路由控制等关键技术。移动 Ad hoc 网络的特点要求其路由协议引入的延时以及维护路由的开销尽量小，同时能够适应网络拓扑结构的动态变化，提供高效的最佳路由，在满足上述条件的基础上，还应尽可能能够提供 QoS 保证。

目前已提出了很多用于 Ad hoc 网络的路由协议，如 DSR、DSDV、AODV 等。其中，按需路由协议 AODV 基本上适合网络拓扑快速变化情况下的数据传输的要求。但是 AODV 在路由发现的过程中使用泛洪策略，在发现到达目的节点的路由之前，将在一定的 TTL 内广播 RREQ 分组，这将产生巨大的网络开销。已有研究采用根据能量、队列长度等性能指标延迟泛洪的算法以选择最佳路径。研究表明延迟泛洪的算法并未限制泛洪，对路由开销的减小有限，且延迟发送 RREQ 分组将延长路由发现的时间，造成延时性能的降低。针对上述问题，本文提出了使用基于平均延时的泛洪算法的 AODV 改进协议 AODV-D (AODV Based on Delay)，它根据节点相邻链路的延时大小，控制泛洪的频率，以实现低延时路径的选择，从而提高网络的整体性能。

1 AODV路由协议

AODV^[4,5]是专为移动 Ad hoc 网络设计的一种路由协议，综合使用了 DSDV 中目标序列号和 DSR 中按需路由发现技术。AODV 路由算法属于按需路由算法，即仅当有源节点需要向某目的节点通信时，才在节点间建立路由，路由信息具有一定的生命期。在实现上，AODV 包括三大部分：路由发现、路由维护和路由删除^[6]。

1.1 路由发现

当一个节点需要和某个节点通信时，如果路由存在且有效，则根据路由发送数据；如果没有到此目的节点的有效路径，它就广播路由请求 RREQ。当其他节点收到 RREQ 时，首先丢弃重复接收的 RREQ，而后根据 RREQ 中的信息建立本节点到源节点的反向路由。一旦 RREQ 分组到达目的节点或者存有目的节点有效路由的中间节点，该节点就发送路由应答 RREP 沿反向路由回传给源节点，否则继续泛洪过程。

中间节点和源节点在收到 RREP 后会建立到目的节点的路由，并更新序列号等信息。源节点收到回传的 RREP 分组，则完成到目的节点路由的建立，源节点开始传输数据。

1.2 路由维护

AODV 协议设计了两种路由维护方式：源节点路由重建和本地修复。

路由维护探测网络拓扑结构是否发生变化，从而决定路径是否可用。当探测到路径中断时，源节点路由重建过程使用 RERR 分组通知所有受影响的源节点，并由源节点重新发起路由发现过程。

另外一种处理方式是由中间节点直接修复路由。检测到链路中断时，中间节点缓存来自源节点的数据流，同时发送路由请求，目的节点如果收到该请求，则回复路由应答，路由修复成功；否则，在一定时间内没有收到路由应答，则仍然启用源节点路由重建。

2 基于延时的泛洪算法

标准 AODV 协议以最小跳数作为路由度量, 由最先收到的 RREP 确定到达目的节点的路由。随着网络负荷的增加与网络规模的扩大, 这种方式容易引起网络中部分节点较其它节点更多地参与通信, 从而发生拥塞或者节点能量耗尽等情况, 影响网络的性能。考虑到上述问题, 本文为 AODV 协议引入了延时度量, 使用基于延时的泛洪算法实现负载均衡与低延时路径选择。

延时是反映链路的好坏的基本指标。设 R_i 代表第 i 个分组的接受时刻, S_i 代表第 i 个分组的时间戳, 即发送时刻, 则第 i 个分组的延时 D_i 为:

$$D_i = R_i - S_i$$

为测定与某一节点相连的链路的延时性能, 修改 hello 分组头与邻居表, 使其能够测量并记录两个邻居节点间链路的延时。在邻居表中添加延时项, 当周期性广播 hello 消息时, 根据发送接收 hello 分组的延时, 更新邻居表中的延时值, 设第 m 条链路收到第 n 个 hello 分组时的链路平均延时为 $Dlink_m$, 根据以下公式计算该链路上的平均延时:

$$Dlink_m = \frac{\sum_{i=1}^n D_i}{n}$$

节点 k 根据邻居表中各个链路的平均延时, 计算连接节点 k 的所有 M 个链路的平均延时 $Dtotal_k$:

$$Dtotal_k = \sum_{m=1}^M \frac{Dlink_m}{M}$$

该值能够反映经过该节点到下一节点的所有路径的平均延时水平。

基于 $Dtotal_k$, 采用如下的泛洪算法: 在路由发现过程中, 中间节点 k 根据自身的平均延时 $Dtotal_k$ 调整转发 RREQ 分组的频率, 当 $Dtotal_k$ 不大于 $delay_threshold$ 时, 使用原 AODV 的泛洪策略, 当 $Dtotal_k$ 大于 $delay_threshold$ 时, 随 $Dtotal_k$ 增大控制 RREQ 分组的转发频率。

设频率调节因子为 α , 转发 RREQ 分组的概率 ρ 可表示为:

$$\rho = \begin{cases} 1, Dtotal_k \leq delay_threshold \\ \frac{1}{\alpha \times Dtotal_k}, Dtotal_k > delay_threshold \end{cases}$$

在如图 2 所示的拓扑结构中, 节点 A1、B1、C1 的 $Dtotal$ 值较大, A2、B2、C2 的 $Dtotal$ 值较小, 源节点 S 要向目的节点 D 发送数据而没有到达 D 的有效路由, 则 S 发起路由发现过程, 向 A1、A2 等节点广播 RREQ。根据上述算法, 节点 A1 的以较低的频率转发 RREQ, A2 则以较高频率转发 RREQ, 其他节点类同, 因此 RREQ 分组沿路径 A2-B2-C2 到达 D 的概率比沿 A1-B1-C1 或其他路径到达 D 的概率高, 这样可以选择了一条延时较小路径。而当延时性能发生变化时, 将选择不同节点传送数据, 从而实现负载均衡。

调整不同的门限 $delay_threshold$ 以及频率调节因子 α , AODV-D 能获得不同的性能。通过实验, 发现 $delay_threshold=0.01$, $\alpha=200$, 协议能获得较好的性能, 以下实验都以该值进行仿真。

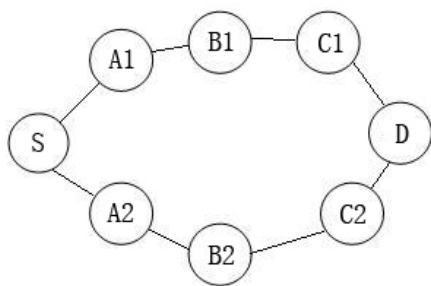


图 1 特定拓扑结构中的路径选择

3 性能评估

3.1 仿真环境

本实验使用 NS2.33 进行仿真，在 1000m x 1000m 的空间中随机配置 50 个节点^[7]，每个节点通信距离为 250m，仿真时间为 300s。节点的运动符合 Random Waypoint 模型。为对比不同网络负载下的协议性能，考虑分组发送率为 1.0、2.0、3.0、4.0、5.0、6.0 分组/s 的流量场景进行比较。详细的模拟参数如表如表 1、表 2 所示：

表 1 节点运动场景参数

参数变量	参数值
节点移动范围(m)	1000 x 1000
仿真时间(s)	600
节点数	50
节点最大速度(m/s)	20
节点停留时间(s)	0

表 2 网络流量场景参数

参数变量	参数值
仿真时间(s)	600
节点通信范围(m)	250
带宽(Mb)	2
业务类型	CBR
最大连接数	20
分组发送率(分组/s)	1.0、2.0、3.0、4.0、5.0、6.0

3.2 仿真结果分析

分组投递率与延时性能的比较如图 2 与图 3 所示，在低负载时，AODV-D 与 AODV 有相近的分组投递率与端到端延时，这是因为在未超过延时门限的情况下，AODV-D 仍采用 AODV 的泛洪策略。在高负载情况下，由于发生拥塞的可能性增大，AODV 的分组投递率在迅速下降，而延时显著增加，AODV-D 能够根据延时选择路由，可以避开拥塞的路径，达到负载均衡的目的，因此在分组投递率与延时方面的性能均优于原协议。

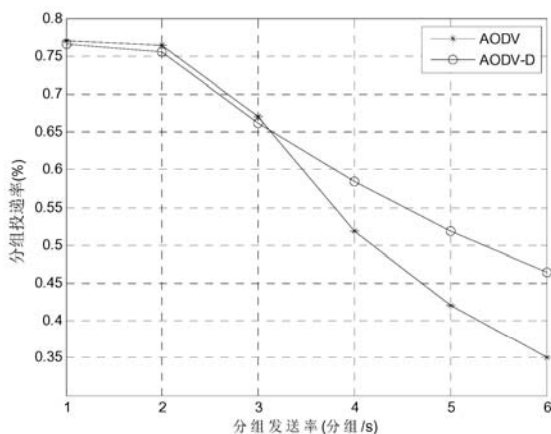


图 2 不同负载下的分组投递率比较

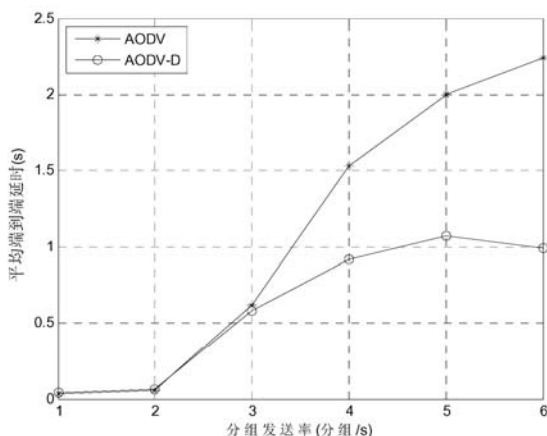


图 3 不同负载下的平均端到端延时比较

路由负荷比较如图 4 所示。与投递率及延时相似，AODV-D 在低网络负载下的路由负荷同 AODV 相当，在高负载下路由负荷均比 AODV 要小的多，且随着网络负载的增大而减小路由负荷。这是因为路由负荷很大程度上是 RREQ 泛洪引起的，AODV-D 限制了 RREQ 泛洪频率，且延时越大泛洪的概率就越小。

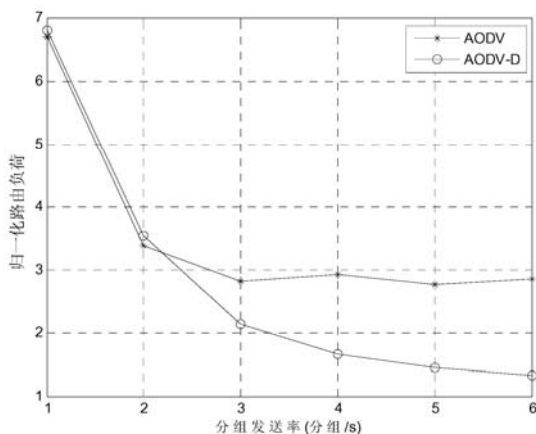


图 4 不同负载下的路由负荷比较

4 结论

传统路由协议在 Ad Hoc 网络中存在很多问题，因而提出了多种路由协议来改进。本文提出基于延时的 AODV-D 协议，NS2 仿真表明，在低负载下 AODV-D 能达到与 AODV 协议相近的性能，同时在高负载下各方面性能都大大优于 AODV，网络效率得到了提高。

参 考 文 献

- [1] S. CORSON, J. MACKER, "Mobile Ad hoc Networking (MANET)," Internet Draft, Oct. 1998.
- [2] ROYER E, TOH CK. A Review of Current Routing Protocols for Ad Hoc Mobile Wireless Networks[J]. IEEE Personal Communications Magazine, 1999, 6(2):46~55
- [3] RAM RAMANATHAN and JASON REDI. A Brief Overview of Ad Hoc Networks: Challenges and Directions. IEEE Comm. Magazine 50th Anniversary Commemorative Issue, 2002(5), 20~22
- [4] CHARLES E. PERKINS and ELIZABETH M. ROYER. "Ad hoc On-demand Distance Vector Routing." Proceedings of the 2nd IEEE Workshop on Mobile Computing Systems and Applications, New Orleans, LA, pp. 90-100, February 1999.
- [5] RFC3561, "Ad hoc On-demand Distance Vector (AODV) Routing." July 2003.
- [6] 于宏毅等.无线移动自组织网[M].北京：人民邮电出版社，2005:176~182.
- [7] 徐雷鸣, 庞博, 赵耀. NS-与网络模拟[M]. 北京：人民邮电出版社,2003: 104~107

作者简介：

陈哲愚(1985-), 男, 福建福州人, 南京邮电大学硕士研究生, 主要研究方向为无线 Mesh 网络路由协议、无线网络跨层设计等。

一种改进的BP量子LDPC译码算法

段 灿 朱修利

(南京邮电大学 通信与信息工程学院, 江苏 南京 210003)

摘要: 稳定子码是目前量子纠错编码技术应用最普遍的技术, CSS 量子码最常用置信传播算法 (BP 算法) 译码。由于量子 LDPC 码固有的简并问题, 当稳定子码中的稳定子生成元具有对称性时, 常规 BP 译码无法译码, 本文研究了稳定子码产生简并问题的根本原因, 并提出了两种解决该问题的有效方法, 即冻结法和随机扰乱法, 打破 BP 算法中置信的对称性使其成功译码, 通过数值计算和常规 BP 译码算法的性能进行比较, 验证了改进 BP 译码算法的可行性以及优越性。

关键词: 置信传播译码; 稳定子码; 冻结法; 随机扰乱法

An improved BP quantum LDPC decoding algorithm

Duan Can Zhu Xiu li

(School of Telecommunication & Information Engineering, Nanjing University of Posts & Telecommunication, Jiangsu Nanjing 210003, china)

Abstract: The Stabilizer codes are the basic fashion of quantum LDPC codes. CSS quantum always accomplish decoding with Belief Propagation algorithm (BP algorithm) technique. As quantum LDPC codes have inherent degeneracy error problem, When the stability generator of Stabilizer codes are exactly symmetrical, the conventional BP algorithm can not decode. In this thesis, the reasons caused of the problem are discussed and two solutions are proposed to solve the problem: freezing and random perturbation. By numerical simulation, the results are analyzed and compared with the result with the conventional BP algorithm, experiments verify the feasibility and superiority of this method.

Keywords: belief propagation algorithm; stablizer codes; Random perturbation

1 引言

量子纠错编码是量子通信和量子计算实用化的基础, 迄今为止, 量子纠错技术日趋完善, 几乎所有经典纠错编码方案都已经被移植到量子领域中。^[1]低密度奇偶校验码 (LDPC) 码以其低复杂度的迭代译码算法和可逼近信道容量限的特性已成为经典通信中最佳的编码技术之一。^[2,3]目前介绍的量子 LDPC 纠错码都是基于 CSS 量子纠错码构造的。^[4,5,6]主要译码方法为 BP 算法。本文主要研究一种改进的 BP 译码算法, 打破置信的对称性而成功译码, 使之更好应用于量子译码环境。

2 量子纠错码的简并

假设有 $E|\psi\rangle$ 为混合量子态，其中 $E \in P_n$ ，设 $E' = ES_c$ ，其中 S 为一算子，则根据量子态的定义， $E|\psi\rangle = E'|\psi\rangle$ ，而且，由于 E, E' 只是由算子 S 区分的，而 S 是由生成元 S_c 组成的，所以 E, E' 具有相同的指错子，这样的话错误 E 和 E' 就无法分辨出来。然而在码字空间中任何一个可以纠正 E 错误的恢复算子 R 也同样可以纠正 E' 的错误，如公式：

$$RE|\psi\rangle = |\psi\rangle \rightarrow RE'|\psi\rangle = RES|\psi\rangle = RE|\psi\rangle = |\psi\rangle \quad (1)$$

简言之，与算子 S 相关的错误不能也不必使用指错子去辨别。这个就是量子纠错码简并的主要特征^[7]。

由于 $\{S_c, T_c, \bar{X}_j, \bar{Z}_j\}$ 是 P_n 的一组标准型矩阵形式，我们可以将错误展开成唯一表现形式，比如 $E = s(E) * t(E) * l(E)$ ，特别注意 $t(E)$ 是分解形式中唯一可以与稳定子生成元反对易的因子。 $\langle T_c \rangle$ 中的每一个因子与 S_c 有独立的联系，于是具有独立互异的指错子。这样的话，当指错子 s 确定时， $t(E)$ 也会被完全确定的，如下式：

$$t(s) = \prod_{c=1}^m T_c^{\frac{1+s_c}{2}} \quad (2)$$

从另一方面来说，由于分解出来的因子式只是作用于 S_c 的，所以不需要计算 $s(E)$ 的具体数值。这样唯一需要关心的算子就是 $l(E)$ ，它的条件概率如下式：

$$p(L|s) \propto \sum_{E:l(E)=L} P(E) = \sum_{S \in S} p(E = St(s)L) \quad (3)$$

最优译码方法必须要能根据 s 得到 L 的最大似然值：

$$L^*(s) = \arg \max_L \{p(L|s)\} \quad (4)$$

由于量子纠错码中简并特性的存在，所以对许多错误的纠错可以使用相同的方法，而不需要每次都使用指错子去分辨错误，从这个程度上说简并可以提高码字特性。为了充分利用这一特性，正确的统计出陪集（具有最大似然可能性的陪集）的概率是十分重要的。然而 BP 译码算法主要是利用对期望 E_q 的计算迭代，完全不考虑码字的陪集结构。这样在使用置信传播译码方法比如 BP 译码算法的过程中，简并将会造成额外的复杂问题。而这点主要是由于码字中结构的对称性，进而影响到译码特性造成的。

假设一个特殊的简单量子码是将一个 0 量子比特编码成两个量子比特。那么它具有两个稳定子生成元 XX 和 ZZ ，并且码字空间 $C = \{(|00\rangle + |11\rangle) / \sqrt{2}\}$ 。假设这些量子比特是通过退极化信道传播的，并且错误 IX 是允许存在的。根据测算，这两个稳定子生成元的指错子为 $s=(+1, -1)$ ，这些错误可以通过 IX, XI, YZ, ZY 中任一种运算过程进行纠正。由于码字的对称性，两个量子比特的边缘概率密度是一样的， $P_1(E|s) = P_2(E|s)$ 。由于没有任何一个纠错方法具有如此的对称性，所以任何基于量子比特概率的译码方法对如此简单的量子码都将失效。

3 针对稳定子码简并问题改进译码算法及仿真

无论在何种情况下，BP 算法的总体结构是不会被改变的。每次迭代过程中，新的置信度产生，根据最大可信度改变的期望 E^{BP} 也重新被估计。如果 E^{BP} 能恢复成一般指错子，程序停止并输出 E^{BP} 。反之，程序将继续进行直到达到预设的不满足停止条件下的迭代次数 T ，这时将会引入一种扰乱机制以确保打破上文所述的对称性。解决该问题的不同方法差异主要体现在打破对称性上选择的差异性上。Tanner 图中短环的存在是十分关键的，它可以使引入的扰乱机制迅速传播至所有具有简并特性错误的量子比特中。本文主要研究了用冻结法和随机扰乱法来实现。

3.1 冻结法

该方法核心思想为在错误检测迭代过程中随机选择其中一个量子比特并将其“冻结”，以此得名。首先我们找到一个不能正常停止的检测 c ， $S_c \cdot E^{\text{BP}} \neq s_c$ ，从中随机选择量子比特 q ，并将其先验概率冻结为 $p_q(E_q) = \delta_{E_q, I}$ 。如果在此过程中，迭代停止了，则可得到先验概率；如果 c 检测过程还是错误的，我们恢复 q 的概率并冻结 c 检测中的量子比特 q' 。如果 c 检测过程是正确的，但是不满足迭代停止的条件，我们选择另一个错误的检测过程 c' 并冻结其中一个量子比特，重复以上步骤，直到满足条件为止。

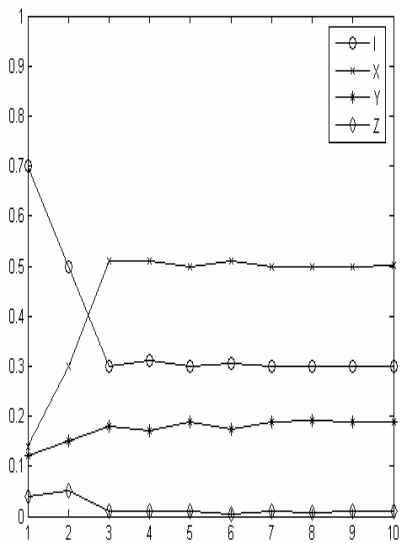


图 1 存在简并问题稳定子码量子比特的置信度

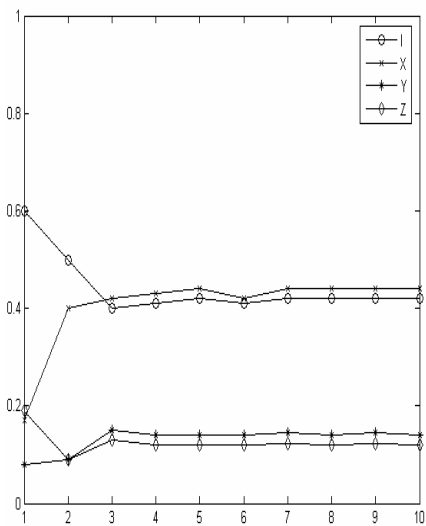


图 2 使用冻结法后稳定子码量子比特的置信度

我们对上节提出的假设进行仿真，引入量子纠错码具有稳定子生成元 XX 和 ZZ ，指错子为 $(+1, -1)$ ，错误模型为退极化信道。其中信息在信道传输中比特错误概率为 P ，即发生比特翻转错误、相位翻转错误以及比特和相位均发生错误的概率各为 $P/3$ 。将冻结法引入本例中，随机冻结技术就会使简并问题迎刃而解。假设我们使 $p_2(E_2) = \delta_{E_2, I}$ ，在一次迭代后， $b_1(E_1) = \delta_{E_1, X}$ ， $b_2(E_2) = \delta_{E_2, I}$ ，这样纠错后结果 $E^{\text{BP}} = XI$ ，即是正确的译出了结果，说明冻结

法针对具有简并问题的量子码可以成功译码，但是根据第一张图发现 I 和 X 的置信度差异性并不明显，说明这种那个方法的纠错强度还有所欠缺。

3.2 随机扰乱法

该方法即是对错误检测过程中量子比特的先验概率进行随机扰乱，我们找到错误检测过程 c ,对 c 中的每个量子比特 q 我们引入扰乱

$$p_q(I_q) \rightarrow p_q(I_q) \tag{5}$$

$$p_q(X_q) \rightarrow (1+\delta_X)p_q(X_q) \tag{6}$$

$$p_q(Y_q) \rightarrow (1+\delta_Y)p_q(Y_q) \tag{7}$$

$$p_q(Z_q) \rightarrow (1+\delta_Z)p_q(Z_q) \tag{8}$$

这里 $\delta_X, \delta_Y, \delta_Z$ 是 $[0, \delta]$ 随机分布变量， δ 是确定值。引入这种扰乱的目的在于在量子比特中建立一种非对称性，从而打破死循环。这种扰乱必须是随机分配的，不然每次错误检测过程中的先验概率都是按相同值递增的，简并问题依然存在。

将随机扰乱法模型引入到上节的示例中，得到了我们预期得到的结果。 $E^{BP} = XI$ ， $p_1(X)$ $p_2(X)$ 具有极大的不对称性。取 $\delta=1$ 对其进行仿真，译码结果都 XI ，这说明使用随机扰乱法后能成功译出具有对称性的稳定子码，即是说能有效解决简并问难，但是当选择的扰乱算子过小时，无法正确译码，说明是存在微扰的，当选择的扰乱算子较大时，稳定子置信度之间差异性也稍大，易于得到正确结果，所以在一定范围数值范围内，可以选择较大的扰乱算子。

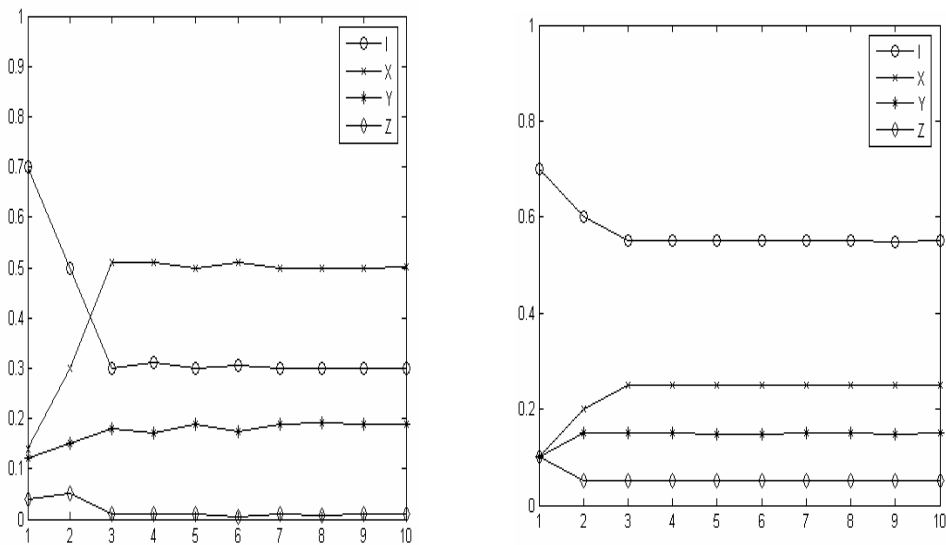


图3 使用随机扰乱法 ($\delta=1$) 的稳定子译码量子比特置信度

4 常规BP译码算法与改进译码算法性能比较

本文仿真了码率都为 $1/4$ 的 LDPC 码，能编 4 个量子比特，编码长度为 16，错误模型为退极化信道，其中信息在信道传输中比特错误概率为 P ，即发生比特翻转错误、相位翻转错误以及比特和相位均发生错误的概率各为 $P/3$ ，迭代上限次数为 1000，纵坐标为误帧率，横坐

标为比特错误的概率。随机扰乱法中 δ 的取值为1。数值仿真表明随机扰乱法以及冻结法在具有简并错误的量子码上译码的可行性，而且，由于随机扰乱算法以及冻结法打破了量子LDPC码的BP译码算法所固有的对称特性，所获得的误帧率性能明显优于常规BP译码算法。

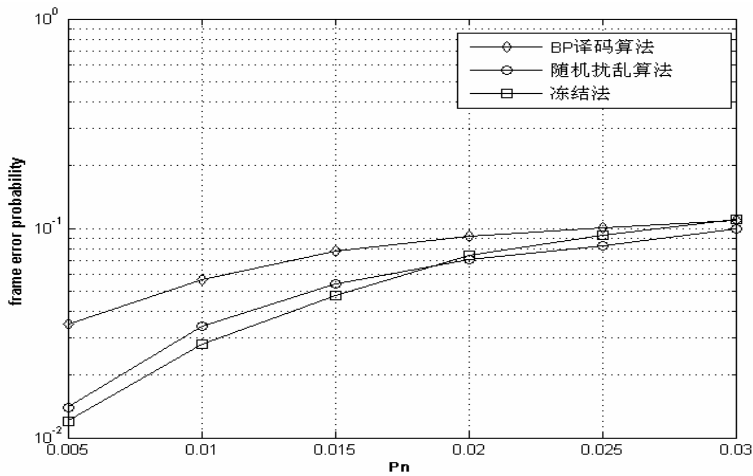


图4 对量子LDPC码(16,4)码三种算法的译码性能比较

5 结束语

本文介绍了量子纠错码中的简并问题，分析简并问题产生的根本原因，并针对稳定子码的简并问题提出了改进方法，对其仿真与常规BP译码算法进行比较，验证了方法的可行性以及优越性。

参 考 文 献

[1] A.M.Steane, Error Correcting Code in Quantum Theory[J], Phys.Rev.Letters, 1996,77(5):pp:793-797

[2] R.Gallager, Low-Density Parity-Check Codes[J],IEEE Trans on Information Theory ,January 1962,pp:21-28

[3] D.J.C.Mackay and R.N.Neal, Near Shannon Limit Performance of Low Density Parity Check Codes[J],IEEE Electronic Letters, Aug 1996,32:pp:1645-1646

[4] A.M.Steane,Quantum Computing and Error Correction[EB/OL], IOS Press, 2001,pp:284-298.Also arXiv/quant-ph/0304016

[5] D.J.C.Mackay, Good error correcting codes based on very sparse matrices, IEEE Trans on Information Theory ,Mae 1999,vol.45:pp:399-431

[6] D.J.C.Mackay ,G.J.Mitchison and P.L.McFadden, SparseGraph Codes for Quantum Error-Correction [EB/OL],2003, arXiv/quant-ph/0304161

作者简介：

段灿(1982—)，女，河北辛集人，南京邮电大学通信与信息工程学院信号与信息处理专业，硕士，主要研究方向为量子LDPC译码方法；

朱修利(1984—)，男，江苏宿迁人，南京邮电大学通信与信息工程学院信号与信息处理专业，硕士，主要研究方向为量子LDPC编码方法。

基于机会式网络编码的两用户协作

李世唐^{1,2} 郑宝玉¹

(1 南京邮电大学 信号处理与传输研究院 南京 210003;

2 福建师范大学 网络安全与密码技术重点实验室 福州 350007)

摘要: 该文在两用户相互协作向基站传送数据的无线通信环境下,研究了基于机会式网络编码的协作策略。在假设所有的信道都存在衰落现象下,比较了非协作传输、传统的协作传输和机会式网络编码的协作传输策略对系统中断概率的影响。仿真结果说明了机会式网络编码协作策略具有较好的性能。

关键词: 网络编码; 用户协作; 存储转发; 系统中断概率

Two Users Cooperation Based on Opportunistic Network Coding

Li Shi tang^{1,2} Zheng Bao yu¹

(1 Institute of Signal Processing and Transmission, Nanjing University of Posts & telecommunications, Nanjing 210003, China; 2 Key Laboratory of Network Security and Cryptology, Fujian Normal University, Fuzhou 350007, China)

Abstract: This paper investigates the strategies of users cooperation based on opportunistic network coding over wireless communication context with two users sending data to the base station. The effects on probability of system outage are compared among non-cooperation, decode-forward cooperation and opportunistic network coding cooperation supposing that all channels exit fading phenomenon. Results of numerical analysis indicate that opportunistic network coding cooperation is better than traditional users cooperation.

Keywords: Network coding; users cooperaton, store and forward, probability of system outage

1 引言

无线信道的多径衰落特性一直以来都是阻碍信道容量增加和服务质量改善的主要原因之一。基于空间分集思想的 MIMO 技术在抗信道多径衰落方面具有明显的优势,并可以在不增加功率和频谱的前提下大幅度提高信道容量,已逐渐被新一代无线通信的主流协议所采纳。然而 MIMO 技术仍然存在很多问题,尤其是小型无线移动终端较难安置多天线的难题使理想

的 MIMO 技术在走向实用的过程中步履维艰。为此, Sendonaris^[1,2]等人首先提出了协作分集的概念, 并分析了协作分集、网络容量和中断概率之间的关系; Laneman^[3]等人深入讨论了协作分集的实现策略, 并给出了分布式空时编码的若干实现策略; Hunter^[4]等人提出了基于编码协作的分集策略。

上述提出的协作分集策略中, 协作策略大都为复制重传、放大重传或空时编码重传等, 都没有利用到业已证明能极大提高无线通信网络性能的新的编码模式—网络编码。网络编码是 Yeung^[5]等人在研究有线组播网络容量时提出来的概念; 他们证明了在中间节点运用网络编码比传统的运用存储转发机制更具优势。随后, Li^[6]等人给出了单信源无噪声有线组播通信网络环境下, 达到组播容量 (multicast capacity) 的线性网络编码的简单代数构造, 为网络编码的实际应用铺平了道路。对于单信源及多信源无噪声有线组播网络, Koetter^[7]等人运用近世代数方法给出了多项式时间复杂度线性网络编码存在性的证明, 同时他们也证明了网络编码可用来解决非遍历的链路故障问题。

网络编码的核心思想是允许中继节点对来自不同信道的数据进行混合 (编码) 后再转发出去, 以获取更高的网络吞吐量、改善链路负载均衡、减少网络节点的能量消耗、提高网络链接的鲁棒性和纠错效率等。由于网络编码的初衷是权衡中继节点的编码能力来提高系统性能, 它本身就包含了协作的思想, 因而将它与协作分集结合起来以对抗无线信道固有的衰落特性是非常合情理的。也正因为这样, 人们已开始把网络编码用于协作分集的研究上: Xiao^[8]等人利用纠错码的代数叠加方法将网络编码运用到协作分集中, 获得了相当大的编码增益; Bao^[9]等人在多用户协作中, 利用编码图动态地匹配网络图, 提出了自适应网络编码协作技术; 在分布式天线系统 (DAS) 和多用户协作系统环境下, Chen^[10]等人研究了网络编码策略实现的分集性能; 在多单播多中继的无线通信模型下, Peng^[11]等人研究了中继节点采用线性网络编码时系统的中断性能和协作分集增益; Chen^[12]等人将网络编码运用到数据交换的通信环境中, 研究其对系统性能的影响; Lu^[13]等人考虑了有多个中继节点的两用户数据交换通信系统, 提出了机会式网络编码策略; Yu^[14]等人比较了译码转发协作与网络编码协作的性能优劣。

在两用户协作向基站发送数据、用户间的信道也存在固有的衰弱现象的通信环境下, 本文研究机会式网络编码协作策略对系统中断性能的影响。与本文密切相关的是文献[8,14], Xiao 等人考虑的情形是: 每个指定发送数据的用户节点把接收到的伙伴信息与自己将要发送的原始信息编码后进行串接再转发出去, 基站通过迭代译码获得两用户发来的数据; 而 Yu 考虑的情形是: 每个指定发送数据的用户节点把接收到的伙伴信息与自己将要发送的原始信息进行串接后再用系统码编码后转发出去, 基站只要能正确译码就能获得两用户发来的数据。本文考虑的情形是: 指定发送数据的用户节点把接收到的伙伴信息与自己将要发送的原始信息进行异或 (XOR) 相加后再编码转发出去; 或者把两组信息编码后再异或相加, 然后再发送出去, 基站根据收到的四个时隙的数据进行译码以获得两用户发来的数据。

本文剩下的部分安排如下: 第 2 节给出系统模型; 第 3 节推导出系统中断概率的表达式; 第 4 节通过实验仿真比较所提协作策略的优劣; 第 5 节对本文进行总结。

2 系统模型

考虑两个用户 A 和 B 把数据发送到基站 D 这样一个无线通信模型。假设初始信号和协作中继信号通过正交时分信道发射, 且用户到基站、用户到用户的信道是相互独立的。

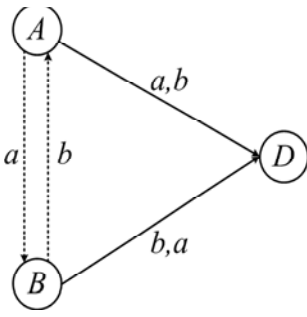


图 1 传统的译码转发协作

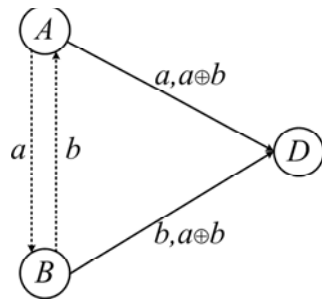


图 2 机会式网络编码协作

为提高分集增益，传统的译码转发协作如下：在第一个时隙，用户 A 广播数据 a ，此时伙伴节点 B 和目的节点接收数据 a ；在第二个时隙，用户 B 广播数据 b ，此时伙伴节点 A 和目的节点接收数据 b ；在第三个时隙，如果用户 A 能正确译出数据 b ，则转发数据 b ；如果无法正确译出数据 b ，则再次发送数据 a ；在第四个时隙，如果用户 B 能正确译出数据 a ，则转发数据 a ；如果无法正确译出数据 a ，则再次发送数据 b （如图 1）。

机会式网络编码的用户协作如下：在第一个时隙，用户 A 广播数据 a ，此时伙伴节点 B 和目的节点接收数据 a ；在第二个时隙，用户 B 广播数据 b ，此时伙伴节点 A 和目的节点接收数据 b ；在第三个时隙，如果用户 A 能正确译出数据 b ，则转发数据 $a \oplus b$ ；如果无法正确译出数据 b ，则再次发送数据 a ；在第四个时隙，如果用户 B 能正确译出数据 a ，则转发数据 $a \oplus b$ ；如果无法正确译出数据 a ，则再次发送数据 b 。

与传统的译码转发协作相比，机会式网络编码的用户协作具有更大的优势。这是因为只要用户 A 到用户 B（或用户 B 到用户 A）的信道条件比较好，那么采用机会式网络编码协作时，B（或 A）将转发编码的数据，此时 A 的数据 a 被发送了三次（从信息论角度来看），而 B 的数据 b 至少被发送了两次（或者 B 的数据 b 都被发送了三次，而 A 的数据 a 至少被发送了两次）；而采用传统的用户协作，用户 A 与 B 的数据都被发送了两次（依赖用户间的信道状况：可能同一信道上发送两次，也可能在不同的信道上各发送一次）。

3 系统中断概率

记 P_{xy} 为节点 X 到节点 Y 的上行链路的 BER。下面考虑系统中断概率，也就是基站 D 无法同时正确接收到用户 A 及用户 B 发送的数据的概率。

定理 1：采用传统的译码转发协作的系统中断概率为：

$$P_{\text{out}}^1 = P_{ad} \left((1 - P_{ba} (1 - P_{ad})) (P_{ab} + (1 - P_{ab}) P_{bd}) \right) + P_{bd} \left((1 - P_{ab} (1 - P_{bd})) (P_{ba} + (1 - P_{ba}) P_{ad}) \right) - P_{ad}^2 P_{bd}^2 \quad (1)$$

证明：设用户 A 的中断概率为 P^a ，用户 B 的中断概率为 P^b ，用户 A、B 同时中断的概率为 P^{ab} ，则

$$P^a = P_{ad} \cdot \left(P_{ba} P_{ad} \cdot (P_{ab} + (1 - P_{ab}) P_{bd}) + (1 - P_{ba}) \cdot (P_{ab} + (1 - P_{ab}) P_{bd}) \right) \quad (2)$$

$$P^b = P_{bd} \cdot \left(P_{ab} P_{bd} \cdot (P_{ba} + (1 - P_{ba}) P_{ad}) + (1 - P_{ab}) \cdot (P_{ba} + (1 - P_{ba}) P_{ad}) \right) \quad (3)$$

$$P^{ab} = P_{ad}^2 \cdot P_{bd}^2 \quad (4)$$

从而可得系统中断概率为：

$$\begin{aligned} P_{\text{out}}^1 &= P^a + P^b - P^{ab} \\ &= P_{ad} (1 - P_{ba} (1 - P_{ad})) (P_{ab} + (1 - P_{ab}) P_{bd}) \\ &\quad + P_{bd} (1 - P_{ab} (1 - P_{bd})) (P_{ba} + (1 - P_{ba}) P_{ad}) - P_{ad}^2 P_{bd}^2 \end{aligned}$$

定理 2： 采用机会式网络编码协作的系统中断概率为：

$$\begin{aligned} P_{\text{out}}^2 &= P_{ad} P_{ad} (1 - (1 - P_{bd})^2 (1 - P_{ab})) + P_{ad} (1 - P_{ba}) (1 - P_{ad}) (P_{bd}^2 P_{ab} + P_{bd} (1 - P_{ab})) \\ &\quad + P_{bd} P_{ad} (1 - (1 - P_{ad})^2 (1 - P_{ba})) + P_{bd} (1 - P_{ab}) (1 - P_{bd}) (P_{ad}^2 P_{ba} + P_{ad} (1 - P_{ba})) \\ &\quad - P_{ad} P_{bd} \cdot (P_{ba} P_{ad} P_{ab} P_{bd} + P_{ba} P_{ad} (1 - P_{ab}) + (1 - P_{ba}) P_{ab} P_{bd} + (1 - P_{ab}) (1 - P_{ba})) \end{aligned} \quad (5)$$

证明：设用户 A 的中断概率为 P^a ，用户 B 的中断概率为 P^b ，用户 A、B 同时中断的概率为 P^{ab} ，则有

$$P^a = P_{ad} P_{ad} (1 - (1 - P_{bd})^2 (1 - P_{ab})) + P_{ad} (1 - P_{ba}) (1 - P_{ad}) (P_{bd}^2 P_{ab} + P_{bd} (1 - P_{ab})) \quad (6)$$

$$P^b = P_{bd} P_{ad} (1 - (1 - P_{ad})^2 (1 - P_{ba})) + P_{bd} (1 - P_{ab}) (1 - P_{bd}) (P_{ad}^2 P_{ba} + P_{ad} (1 - P_{ba})) \quad (7)$$

$$P^{ab} = P_{ad} P_{bd} \cdot (P_{ba} P_{ad} P_{ab} P_{bd} + P_{ba} P_{ad} (1 - P_{ab}) + (1 - P_{ba}) P_{ab} P_{bd} + (1 - P_{ab}) (1 - P_{ba})) \quad (8)$$

系统中断概率为： $P_{\text{out}}^2 = P^a + P^b - P^{ab}$ ，将式(6)，(7)与(8)代入即得。

若假设用户 A 或 B 到基站 D 的信道的 BER 均为 p ，而用户间信道的 BER 均为 q ，则上述系统中断概率分别为：

$$P_{\text{out}}^1 = 2p^2 + 2pq - 4p^2q - 2pq^2 + 4p^2q^2 + 2p^3q - 2p^3q^2 - p^4 \quad (9)$$

$$P_{\text{out}}^2 = 2p^2q + 3p^3 - 2p^4 - 3p^3q + p^4q \quad (10)$$

$$P_{\text{out}}^1 - P_{\text{out}}^2 = 2p^2 + 2pq - 6p^2q - 2pq^2 + 3p^2q^2 + 5p^3q - 2p^3q^2 - 3p^3 + 2p^4 - p^4q \quad (11)$$

由上述三个式子可得，当 $p \square q \square 1$ 时，2 种用户协作策略都获得了完全分集阶数 2；又由于 $p = q \square 1$ 时， $P_{\text{out}}^1 - P_{\text{out}}^2 > 0$ ，故基于机会式网络编码的用户协作策略要优于传统的用户协作策略。下面通过数值分析来比较这两种协作策略的优劣。

4 数值分析

考虑每个用户每次发送数据的发射功率都为 E 的情形。假设用户到基站的信道增益服从均值为 0 方差为 1 的复高斯分布： $h \square CN(0,1)$ ，而用户间信道的 BER 都为 q ，并且假设采用 BPSK 调制，则相应的上行链路的 BER 为：

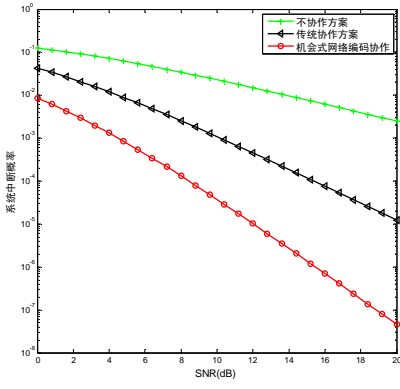
$$p = E \left[Q \left(\sqrt{2|h|^2 \text{SNR}} \right) \right] = \frac{1}{2} \left(1 - \sqrt{\frac{\text{SNR}}{1 + \text{SNR}}} \right) \quad (12)$$

当 $\text{SNR} \square 1$ 时，利用泰勒展开可得：

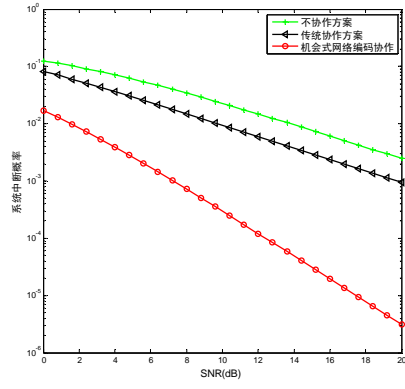
$$p \approx \frac{1}{4 \cdot \text{SNR}} = \frac{N_0}{4E} \quad (13)$$

其中 N_0 为单边高斯噪声功率。

下面分别给出 $q = 0$ ， $q = 0.25$ ， $q = 0.5$ 及 $q = 0.75$ 时系统中断概率与 SNR 的关系图。

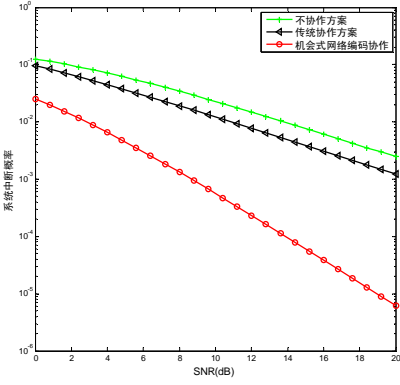


(a) $q = 0$

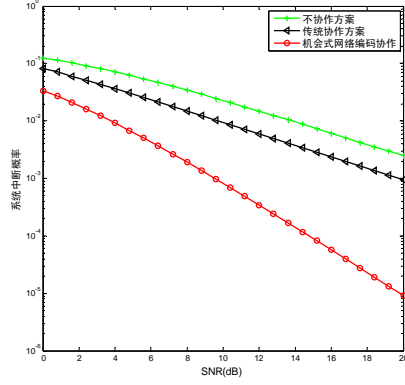


(b) $q = 0.25$

图 3 系统中断概率比较



(a) $q = 0.5$



(b) $q = 0.75$

图 4 系统中断概率比较

从仿真图（图 3，图 4）可以看出，两种协作策略都明显优于不协作的策略。但随着 q 的逐渐增大，传统的协作策略与不协作策略之间的系统中断概率曲线之间的间距变化不明显，而机会式网络编码的协作策略与传统的协作策略的系统中断概率曲线之间的间距逐渐缩小。当 $q=1$ 时，也就是每个用户都无法正确获得伙伴用户广播的信息，这时机会式网络编码的协作策略与传统的协作策略的系统中断概率曲线相互重合。此外，由式 (11) 可以看出，当 p 与 q 满足关系式：

$$2p - 6p^2 + 5p^3 - p^4 = (4p - 6p^2 + 4p^3)q \quad (14)$$

时，机会式网络编码的协作策略与传统的协作策略系统中断概率的差距达到最大。但是，上述两种协作传输的优越性的代价是每个用户需要消耗更多的能量，因而需要考虑系统中断概率与能量效率的折衷关系，这也是我们未来的研究方向之一。

5 结论

在两用户相互协作向基站传送数据的无线通信环境下，该文研究了基于机会式网络编码

的协作策略。在假设所有的信道都存在衰落现象下，推倒出了传统的协作传输和机会式网络编码的协作传输策略下的系统中断概率。仿真结果说明：（1）传统的协作传输和机会式网络编码的协作传输策略都要优于不协作的传输策略；（2）当用户间的信道状况很差时，机会式网络编码的协作传输策略的有时并不明显；（3）机会式网络编码的协作传输策略始终优于传统的协作传输。

参 考 文 献

- [1] Sendonaris A, Erkip E, and Zhang B. User cooperation diversity-part I: System description[J]. IEEE Transactions on Communications, 2003, 51(11):1927-1938.
- [2] Sendonaris A, Erkip E, and Zhang B. User cooperation diversity-part II: Implement Aspects and Performance Analysis[J]. IEEE Transactions on Communications, 2003, 51(11):1939-1948.
- [3] Laneman N, Tse D N C, and Wornell G W. Cooperative diversity in wireless networks: Efficient protocols and outage behavior[J]. IEEE Transactions on Information Theory, 2004, 50(12):3062-3080.
- [4] Nosratinia A, Hunter T E, and Hedayat A. Cooperative communication in wireless network[J]. IEEE Communications Magazine, 2004, 42(10):74-80.
- [5] Ahlswede R, Cai N, Li S Y R, et al.. Network information flow[J]. IEEE Transactions on Information Theory, 2000, 46(4):1204-1216.
- [6] Li S Y R, Yeung R W, and Cai N. Linear network coding[J]. IEEE Transactions on Information Theory, 2003, 49(2):371-381.
- [7] Koetter R and Médard M. An algebraic approach to network coding[J]. IEEE/ACM Transactions on Networking, 2003, 11(5):782-795.
- [8] Xiao L, Fuja T E, Klierer J, et al.. A network coding approach to cooperative diversity[J]. IEEE Transactions on Information Theory, 2007, 53(10):3714-3722.
- [9] Bao X and Li J. Adaptive network coded cooperation (ANCC) for wireless relay networks: Matching code-on-graph with network-on-graph[J]. IEEE Transactions on Wireless Communication, 2008, 7(2):574-583.
- [10] Chen Y, Kishore S, and Li J. Wireless diversity through network coding[C]. in Proc. IEEE Wireless Communications and Networking Conference (WCNC), Las Vegas, USA. 2006, 3:1681-1686.
- [11] Peng C, Zhang Q, Zhao M, et al.. On the performance analysis of network-coded cooperation in wireless networks[C]. in Proc. IEEE 26th International Conference on Computer Communications (INFOCOM), Ancharage, AK. 2007, 1460-1468.
- [12] Chen W, Letaief K B, and Cao Z. Opportunistic network coding for wireless networks[C]. in Proc. IEEE International Conference on Communications (ICC), Glasgow, Scotland. 2007: 4634-4639.
- [13] Lü L, Yu H, and Yang J. Opportunistic cooperative network-coding based on space-time coding for bi-directional traffic flows[C]. in Proc. IEEE Workshop on Network Coding, Theory, and Applications (NetCod), Hong Kong, China. 2008: 1-6.
- [14] Yu M, Li J, and Blum R. S., User cooperation through network coding[C], in Proc. IEEE International Conference on Communications (ICC), Glasgow, Scotland. 2007: 4064-4069.

一种多无线电多信道环境下新的信道状态预测算法

黄 川^{1,2} 郑宝玉¹

(1 南京邮电大学信号处理与传输研究院 南京 210003;

2 福建师范大学数学与计算机科学学院 福州 350007)

摘 要: 基于部分可测马尔科夫(POMDP)模型, 结合认知无线电频谱侦测技术, 提出了一种新的多无线电多信道环境下信道状态预测算法。该算法通过对信道状态历史信息分析, 推导出信道信念状态(belief state)的初始分布和转移概率, 并以此选择出具有最佳回报的信道以供接入, 从而达到提高信道利用率的目的。仿真结果表明算法性能要优于传统算法。

关键词: 认知无线电; 多无线电多信道; 部分可测马尔科夫模型(POMDP); 频谱侦测

A Novel Channel State Prediction Algorithm under Multi-radio Multi-channel Environment

Huang Chuan^{1,2} Zheng Bao-yu¹

(1 Institute of Signal Processing and Transmission, Nanjing University of Posts and Telecommunications, Nanjing 210003, China; 2 School of Mathematics and Computer Science, Fujian Normal University, Fuzhou 350007, China)

Abstract: Based on the theory of Partially Observable Markov Decision Process (POMDP) model, a novel multi-radio multi-channel channel state prediction algorithm using spectrum sensing techniques of cognitive radio is proposed. By the analysis of the channel state historical information, the initial distribution of the belief state and transition probability is derived and the channel with optimal reward is selected for unlicensed user to improve the spectrum utilization. The simulation results demonstrate that the proposed algorithm has better performance than classical algorithms.

Keywords: cognitive radio; multi-radio multi-channel; Partially Observable Markov Decision Process (POMDP); spectrum sensing

1 引言

新兴的认知无线电(Cognitive Radio)技术为解决非授权用户有效利用闲置频谱, 提高频谱利用率提供了可能^[1]。由于异构网络的未知性, 使得次用户很难获得不同信道的精确状态, 而必须依据当前的不完全状态信息做出决策。因此对异构网络的未知性采用部分可测马尔科夫模型(POMDP)建立信道状态信息模型并以此做出选择最佳信道的决策具有合理性^{[2][3]}。文献[2]提出了一种基于 POMDP 模型的认知 MAC 协议来解决异构网络中信道选择决策问题。文献[3]则在文献[2]的基础上提出更为具体的以 POMDP 模型为基础的频谱检测策略、信道选择决策

以及接入策略的联合优化方案。上述研究仍存在着两个问题：信道状态间的转移概率是事先给定的，无法根据实际情况变化；由于只有一个无线电收发器，使得用户不得不经常中断数据传输过程转而检测授权用户（或称为主用户）是否出现，从而浪费信道带宽。

针对上述问题，本文通过建立以 POMDP 模型为基础的多无线电系统^[4]模型，提出一种与认知无线电频谱侦测技术相结合的在异构网络环境下的新型信道状态预测算法 CSPA (Channel State Prediction Algorithm)。该算法通过建立异构网络信道状态的 POMDP 模型来构造未知信道状态的预测方法，以使得次用户能获得最佳带宽回报，从而达到提高信道利用率的目的。

本文第二节介绍异构网络环境及多无线电系统模型，第三节首先介绍典型的 POMDP 模型，接着具体描述信道预测算法及算法收敛性证明，第四节是算法的仿真和性能分析，第五节是本文的结论。

2 系统模型

假设网络中每个次用户配备两个独立的具有认知无线电功能的无线电收发设备，当次用户从网络 A 向不同的网络 B 移动时，其一个无线电保持与网络 A 的连接，称为工作无线电；而另一个无线电处于认知侦测状态，则称为观测无线电。由观测无线电侦测到而组成的信道状态历史信息由离散观测时间序列 DT 内的一系列行为 a 、回报 r 、观测状态 z 和应答状态 k 的序列 h 组成，有：

$$h = \{(a_1, r_1, z_1, k_1), \dots, (a_t, r_t, z_t, k_t)\}, \quad t \in DT \tag{1}$$

同时假设在网络 B 中存在 $N(0 < N < \infty)$ 个独立授权信道，每个信道在某一个时隙 t 内处于空闲 (1) 或忙 (0) 状态，则网络在某个时隙 t 内的信道状态为 $S(t) = [S_1(t), \dots, S_M(t)]$ ，其中 $M = 2^N$ ，且 $S_i(t) = s_1(t) \cdots s_N(t)$ ， $s_j(t) \in \{0, 1\}$ ，每个信道的带宽表示为 BW_j ， $j = 1, \dots, N$ 。由于环境的未知性，设次用户可侦测到的信道数目 $n \leq N$ 。

3 基于POMDP的信道状态预测算法

3.1 POMDP模型

一个典型 POMDP 模型可用六元组表示为 $\langle S, A, T, R, Z, O \rangle$ ^[5]。本文中 S 为系统中有限信道状态集合； A 为次用户采取的有限行为（观测，接入）的集合，用 $A = \{a_1, a_2\}$ 表示； T 表示当前信道状态 s 在行为 a 的作用下变为 s' 的转移函数，记为 $T(s, a, s')$ ； R 为瞬时回报函数，记为 $R(s, a)$ 表示； Z 为用户对系统状态的有限观测状态集合； O 为观测函数，记为 $O(s', a, z)$ 。此外， $k_i(t) \in \{0, 1\}$ 表示次用户在执行行为 a 后得到的应答（这里设应答是无错的）。

由于 S 是未知的，采用信念状态空间 B 来表示信道状态的概率分布， $B = \{b: \sum_{s \in S} b(s) = 1.0, \forall s \in S, b(s) \geq 0\}$ ，其中 $b(s)$ 表示信道处于状态 s 的概率。根据 Bayes 法则可得在 $t+1$ 时隙信念状态更新的表达式：

$$b_{t+1}(s') \square \tau(b_t, a_t, z_{t+1})(s') = \frac{\sum_{s \in S} T(s, a_t, s') O(s', a_t, z_{t+1}) b_t(s)}{\sum_{s' \in S} O(s', a_t, z_{t+1}) \sum_{s \in S} T(s, a_t, s') b_t(s)} \quad (2)$$

对于任何时隙 t ，信念状态是一个充分统计量^[6]，从而表明不会有任何历史信息可以增加次用户的回报。同时定义次用户得到的瞬时回报为空闲信道的带宽：

$$R_B(b, a) = \sum_{s \in S} b(s) R(s, a) = \sum_{s \in S} \sum_{i=0}^n b(s=1) B W_i \quad (3)$$

策略 p 表示次用户在当前信念状态 b 的情况下使用策略 p 从而选择行为 a 。解决一个 POMDP 问题的目标是找寻一个最佳策略 π' ，以使用户获得的总折扣回报达到最大值，即次用户所能获得的最大信道带宽，用值函数表示：

$$V^*(b) = \max_{a \in A} [R_B(b, a) + \zeta \sum_{z \in Z} \Pr(z | b, a) V^*(\tau(b, a, z))] \quad (4)$$

其中 ζ ($0 < \zeta \leq 1$) 为折扣因子。

3.2 基于POMDP模型的CSPA算法

以 POMDP 模型为基础的多无线电多信道信道状态预测算法 CSAP 分为两个阶段：1) 观测阶段：通过一段时间的观测以获得系统环境信息并记录到 h 中，期间的执行接入行为仅为向相应频段发送探测包，并未真正执行信道接入操作；2) 预测阶段：通过 h ，对信道的初始状态分布，状态转移概率和观测概率进行估计，接着利用启发式算法找出具有最大折扣回报的策略 π' ，在接入时隙次用户按其接入。

在预测阶段，对次用户来说，信道状态个数 M 是指数的，要计算出最大折扣回报是很困难的，但所幸的是实际网络中的信道一般是独立的，则有如下定理。

定理 假设 n 个独立信道，有 $\Lambda = [\lambda_1, \dots, \lambda_n]$ ，其中 λ_i 为信道 i 在某个时隙 t 开始时刻所处的状态，则 L 是信道状态 $S_i(t)$ 的充分统计量。

证明略。

根据以上定理，POMDP 模型中信念状态空间 $B = \{b(S_i(t)), i=1, \dots, n\}$ 可简化为 $B = \{b(s_k(t)), k=1, \dots, n\}$ ，从而信念空间维度由 2^n 降为 n 。则对每个信道 i 来说，其最大折扣回报表达式为：

$$V_i^*(b_i) = \max_{a_i \in A} [R_B(b_i, a_i) + \zeta \sum_{z_i \in Z} \Pr(z_i | b_i, a_i) V_i^*(\tau(b_i, a_i, z_i))] \quad (5)$$

其中 $R_B(b_i, a_i) = b(s_i=1) B W_i$ ，而对次用户有：

$$V^*(b) = \max_{i=1, \dots, n} V_i^*(b_i) \quad (6)$$

对于在时隙 t 内的每一个信道，可设 θ_i 为在时隙 t 内信道 i 状态为空闲的概率，即 $b(s_i(t)=1) = \theta_i$ ，而状态为忙的概率为 $b(s_i(t)=0) = 1 - \theta_i$ 。则在 T 个时隙中，信道 i 状态为空闲出现次数为 γ_i 服从二项式分布 $b(\gamma_i; T, \theta_i)$ ，则据 Bayes 法则，有 θ_i 对 h 的条件概率密度：

$$f(\theta_i | h) = \frac{q(\theta_i) \binom{T}{\gamma_i} \theta_i^{\gamma_i} (1 - \theta_i)^{T - \gamma_i}}{\int_0^1 q(\theta_i) \binom{T}{\gamma_i} \theta_i^{\gamma_i} (1 - \theta_i)^{T - \gamma_i} d\theta_i} \quad (7)$$

其中 $q(\theta_i)$ 为 θ_i 的先验分布。对于次用户来说，信道 i 处于空闲或忙状态是等可能的，故

先验分布 $q(\theta_i)$ 为 $[0,1]$ 上的均匀分布, 则有:

$$f(\theta_i | h) = \frac{\theta_i^{\gamma_i} (1 - \theta_i^{T - \gamma_i})}{\int_0^1 \theta_i^{\gamma_i} (1 - \theta_i^{T - \gamma_i}) d\theta_i} \quad (8)$$

用 θ_i 对 h 的条件期望 $E\{\theta_i | h\}$ 估计信道状态为空闲的概率:

$$b_i(s_i = 1) = E\{\theta_i | h\} = (\gamma_i + 1)/(T + 2) \quad (9)$$

信道状态转移概率 $T(s, a, s')$ 也是未知的。设 p_{ss}^a 为信道 i 执行行为 a 后状态从 s 转移到 s' 的转移概率, 其中 $s, s' \in \{0, 1\}$, 则有向量 $P_i = (p_{ss_k}^a, a \in A, k = 1, \dots, |S|)$ 。在 T 个时隙中信道 i 的状态从 s 到 s' 的转移次数向量 $\phi_i = (\phi_{ss_k}^a, a \in A, k = 1, \dots, |S|)$, 在执行行为 a 的条件下 P_i 服从 *Dirichlet* 分布, 有 $(p_{ss_1}^a, \dots, p_{ss_{|S|}}^a) \sim \text{Dir}(\phi_{ss_1}^a, \dots, \phi_{ss_{|S|}}^a)$, 其期望为 $E(p_i) = \phi_{ss_k}^a / \sum_{k=1}^{|S|} \phi_{ss_k}^a$ 。当信道状态转移后, 向量 $\phi_i' = \phi_i + \delta_{ss}^a$, 其中向量 δ_{ss}^a 为 $|\delta_{ss}^a| = |S|$, 且 $\delta_{ss}^a[s' = j] = 1$, 其余为 0。则用期望值估计转移概率为:

$$T_\phi(s, a, s') = \phi_{ss}^a / \sum_{s' \in S} \phi_{ss}^a \quad (10)$$

在频谱侦测中由于存在漏检和虚警现象, 因此对于次用户来说, 所观测到的信道状态不一定与信道真实状态相符。假设信道为 AWGN, p_d 为检测概率, p_f 为虚警概率, 且采用能量检测器的频谱侦测方法^[2], 则:

$$p_f = \Pr\{Y > \tau | H_0\} = 1 - \Gamma(L/2, \tau/2\sigma_0^2) \quad (11)$$

$$p_d = \Pr\{Y > \tau | H_1\} = 1 - \Gamma(L/2, \eta\sigma_0^2/\sigma_1^2) \quad (12)$$

由于信道真实状态的未知性, 次用户可根据执行行为 a 后得到的应答信息 k 来验证观测状态的正确与否, 有:

$$O(s', a, z) = \Pr(z | s', a) = \Pr(z | k, a) = \begin{cases} p_d & k = z = 0 \\ 1 - p_f & k = z = 1 \\ p_f & k = 1, z = 0 \\ 1 - p_d & k = 0, z = 1 \end{cases} \quad (13)$$

从而可求出次用户可获得的最大折扣回报为:

$$\begin{aligned} V^*(b) = \max_{i=1, \dots, n} \max_{a_i \in A} [& b(s_i = 1)BW_i + \zeta[p_d p_{00}^a b(s_i = 0) + p_d p_{10}^a b(s_i = 1) + p_f p_{01}^a b(s_i = 0) \\ & + p_f p_{11}^a b(s_i = 1)]V_i^*(\tau(b_i, a_i, 0)) + \zeta[(1 - p_d)p_{00}^a b(s_i = 0) + (1 - p_d)p_{10}^a b(s_i = 1) \\ & + (1 - p_f)p_{01}^a b(s_i = 0) + (1 - p_f)p_{11}^a b(s_i = 1)]V_i^*(\tau(b_i, a_i, 1))] \end{aligned} \quad (14)$$

其中信念状态更新为:

$$\tau(b, a, z) = \begin{cases} \frac{[p_{00}^a b(s = 0) + p_{10}^a b(s = 1)]}{p_d[p_{00}^a b(s = 0) + p_{10}^a b(s = 1)] + p_f[p_{01}^a b(s = 0) + p_{11}^a b(s = 1)]} & k = z = 0 \\ \frac{[p_{01}^a b(s = 0) + p_{11}^a b(s = 1)]}{(1 - p_d)[p_{00}^a b(s = 0) + p_{10}^a b(s = 1)] + (1 - p_f)[p_{01}^a b(s = 0) + p_{11}^a b(s = 1)]} & k = z = 1 \\ \frac{p_f[p_{01}^a b(s = 0) + p_{11}^a b(s = 1)]}{p_d[p_{00}^a b(s = 0) + p_{10}^a b(s = 1)] + p_f[p_{01}^a b(s = 0) + p_{11}^a b(s = 1)]} & k = 1, z = 0 \\ \frac{(1 - p_d)[p_{00}^a b(s = 0) + p_{10}^a b(s = 1)]}{(1 - p_d)[p_{00}^a b(s = 0) + p_{10}^a b(s = 1)] + (1 - p_f)[p_{01}^a b(s = 0) + p_{11}^a b(s = 1)]} & k = 0, z = 1 \end{cases} \quad (15)$$

4 仿真结果及分析

为了测试 CSPA 算法的性能,本文引入随机接入算法^[9] (Random Access Algorithm, RAA) 与之比较。在 RAA 算法中,次用户在刚进入未知新网络时,不使用信道状态预测方法,而是通过每一接入时隙开始时刻的侦测来获知可以采用的若干信道,随机选择其中一个接入。为了体现公平性,设新网络中每一信道的带宽均为 1 个单位,且每个时隙为 1 个单位时间。设折扣因子 $\zeta=1$ 。同时比较 CSPA 算法与理想情况下所获得的信道带宽回报,即与不存在漏检和虚警现象条件下的对比。图 3 表明信噪比 SNR 和检测样本数目的不同时虚警率 P_f 值对检测率 P_d 的影响,从图中可知当 $SNR=10, L=5$ 时,虚警率 P_f 的变化对检测率 P_d 的影响最小,且 P_d 值在 0.9 到 1 之间变动。因此设 $P_d=0.95$ 。仿真中算法对 10000 个随机信道进行运算,然后取回报的平均值(单位每时隙)。

图 4 中设观测时隙 $T_O=30$,接入时隙 $T_A=30$,信道数目 $n=2$ 。从图 4(a)可以看到,CSPA 算法在刚开始的时隙获得约 0.55 的平均回报值,在第 14 个接入时隙上升至 0.73,后增长平稳,逐渐趋近于 0.74;RAA 算法每个时隙的平均回报值稳定在 0.5 左右;在理想情况下,平均回报值稳定在 0.755 左右。图 4(b)对两种算法获得的平均回报值进行比较,图中表明 CSPA 算法能获得的平均回报值比 RAA 算法平均多出约 43%。图 4 表明次用户采用 CSPA 算法,在每个时隙都能够取得最佳的接入策略,因此获得的信道带宽平均回报优于采用传统认知无线电随机频谱接入方式,但由于频谱侦测中漏检和虚警现象的存在使得算法与无漏检和虚警现象存在的理想值(Ideal Value)还有一定的偏差。

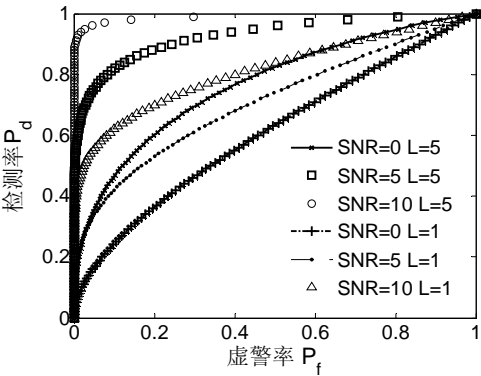


图 3 虚警率和检测率的关系

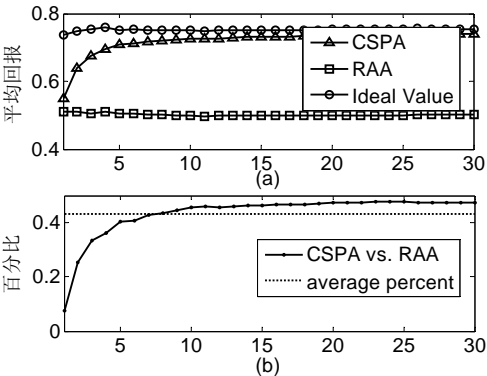


图 4 CSPA 算法与 RAA 算法的对比

5 结论

传统的单无线电技术已无法满足未来异构网络融合的需要,而新兴的多无线电技术为异构网络无缝连接的实现提供了可能。考虑到无线用户节点进入到新网络覆盖区域时,对新网络环境中信道状态的未知性,本文基于 POMDP 模型提出了一种在多无线电多信道环境下带有认知无线电频谱侦测功能的信道状态预测算法(CSPA),以实现用户在多信道切换时能得到最佳信道带宽回报。仿真结果表明 CSPA 算法获得相对于传统认知无线电频谱接入方式更高

的信道带宽，并接近无漏检和虚警现象的理想情况，从而有效地提高了信道利用率。

参 考 文 献

- [1] I. F. Akyildiz, W. Y. Lee, M. C. Vuran, et al. A Survey on Spectrum Management in Cognitive Radio Networks[J]. IEEE Communications Magazine, 2008, 46(4): 40-48.
- [2] Q. Zhao, L. Tong, A. Swami, et.al. Decentralized Cognitive MAC for Opportunistic Spectrum Access in Ad Hoc Networks: A POMDP Framework[J]. IEEE Journal on Selected Areas in Communications, 2007, 25(3): 589-600
- [3] Y.X. Chen, Q. Zhao, A. Swami. Joint design and separation principle for opportunistic spectrum access in the presence of sensing errors[J]. IEEE Trans. on Information Theory, 2008, 54(5): 2053-2071.
- [4] G. Piao, K. David. Multi-Standard Radio Resource Management for Integrated Voice and Data Services[C]. IEEE 65th Vehicular Technology Conference, Dublin, Irish, Apr. 22- 25, 2007, 990-995.
- [5] L. P. Kaelbling, M. L. Littman, A. R. Cassandra. Planning and acting in partially observable stochastic domains[J]. Artificial Intelligence, 1998, 101(1): 99-134.
- [6] R. Smallwood, E. Sondik. The optimal control of partially observable Markov processes over a finite horizon. Operations Research[M], 1971: 1071-1088.

作者简介:

黄川 (1979-), 男, 福建福州人, 南京邮电大学, 在职博士研究生, 主要研究方向为无线通信与网络信号处理等。

郑宝玉 (1945-), 男, 福建闽侯人, 南京邮电大学, 教授、博导, 上海交通大学兼职教授、博导, 主要研究方向为无线通信与网络信号处理、现代通信中的智能信号处理与信息处理等。

协作通信的能量有效性和时间有效性研究

季 薇 郑宝玉 江若宜

(通信与信息工程学院 南京邮电大学, 南京 江苏 210003)

摘 要: 无线通信中的能耗来所有功率放大器的能耗和其他电路能耗。在以短距离传输为主的传感器网络中, 电路能耗成为不可忽略的一部分, 甚至有可能超过功放能耗成为通信能耗的主导。如何通过节点间的协作提高网络的能量有效性成为人们普遍关注的一个热点。本文将自适应调制引入协作通信系统模型, 对于无线传感器网络环境下的能量有效性和时间有效性的折中进行了分析, 并提出了发送功率受限条件下的协作通信系统性能优化策略, 结合网络路由结构对协作通信系统的性能进行了跨层优化。

关键词: 协作分集; 能量有效性; 时间有效性; 无线传感器网络

Energy Efficiency and Time Efficiency in Cooperative Communications

Ji Wei Zheng Bao yu

(College of Telecommunications & Information Engineering, Nanjing University of Posts & Telecommunications, Nanjing, 210003, China)

Abstract: Energy consumption in wireless communication includes amplifier energy consumption and circuit energy consumption. In short-distance communications, circuit energy consumption cannot be omitted, how to exploit the advantage of cooperative diversity in wireless communication becomes a crucial problem. In this paper, adaptive modulation is introduced in cooperative wireless sensor networks. And we consider the tradeoff of energy consumption and time delay in network environment and propose a strategy to optimize above performance with limited transmission power. Cross-layer design is used to achieve this strategy by considering routing architecture

Keywords: Cooperative Diversity; Energy Efficiency; Time Efficiency; Wireless Sensor Networks

1 引言

无线通信中的能耗来自两个方面, 即所有功率放大器的能耗(又称路径能耗)和其他电路能耗。在大距离传输中, 路径能耗占主导地位, 因而人们常常将电路能耗忽略不计。而在

国家自然科学基金项目(60972039); 国家 863 计划项目(2009AA01Z241); 江苏省高校自然科学研究项目(09KJB510012); 江苏省自然科学基金重大项目(BK2007729)。

短距离传输中，通信能耗不再由路径能耗主导，电路能耗成为不可忽略的一部分，甚至有可能超过路径能耗成为通信能耗的主导。如何通过节点间的协作^[1-6]提高网络的能量有效性成为人们普遍关注的一个热点。

协作通信中的能量有效性研究最具代表性的是 S. Cui 和 S. K. Jayaweera 的工作。S. Cui 等人首次将电路能耗纳入到网络通信能耗的计算中，建立了无线传感器网络中的协作通信模型。在这个模型中，作者假设簇内节点间距离远小于簇间传输距离，因而簇内通信能耗可忽略不计，将这个通信过程称为本地通信；簇间通信则被抽象成虚拟 MIMO 通信，将这个通信过程称为远程通信。此基础上，S. Cui 结合 Alamouti 空时编码分析了上述模型的能量消耗，并得出协作通信机制在传输距离较远的情况下能有效降低网络能耗的结论^[7-9]。S. K. Jayaweera 等人在 S. Cui 工作的基础上，提出当协作节点之间的距离不可忽略时，本地通信能耗也将成为网络能耗的重要部分，并结合信道训练开销 (training overhead) 设计了一个新的能耗模型。与 S. Cui 的模型不同的是，在接收端，S. K. Jayaweera 使用了类似于汇聚节点的数据收集节点 (Data Gathering Node, DGN)^[10-12]。然而，上述研究工作仍存在一定的缺陷。S. Cui 的研究中使用 Chernoff 界来估计瑞利衰落信道下的系统差错率，并据此计算系统的能耗；但 Chernoff 界只在高信噪比环境下比较准确，若能够进一步得到更精确的结果，则会更有利于实际系统的设计及性能分析。虽然 S. K. Jayaweera 对瑞利衰落信道下的差错概率进行了初步地统计分析，但结果并不完善；且 S. K. Jayaweera 提出的模型需要对 DGN 进行特殊设置，这在现有的大多数无线传感器网络中并不适用。因而，需要对协作通信中的能量有效性和时间有效性问题进行更为系统和深入的研究。

本文从协作通信的能量有效性角度出发，研究协作式 MIMO 技术在无线传感器网络中的应用，拟解决“是否协作”，“何种情况下协作”以及“如何协作”的问题，并由此引发出对时间有效性的考量。将自适应调制引入协作通信系统模型，以提高协作传输的节能效果，减少协作通信时延。在此基础上，对于网络环境下的能量有效性和时间有效性的折中进行了分析，并提出了发送功率受限条件下的协作通信系统性能优化策略，结合网络路由结构对协作通信系统的性能进行了跨层优化。

2 无线通信基本能耗分析模型

通常，无线链路的功耗主要来自两个方面，即电路功耗 P_c 和功放功耗 P_{PA} 。对固定速率的系统而言，传输一个比特所需要的总能量为

$$E_{bt} = (P_c + P_{PA}) / R_b \quad (1)$$

功放能耗 P_{PA} 取决于发送能耗 P_{out} ，它与信道的衰落有关。当信道只经历路径衰落时， P_{out} 可表示为

$$P_{out} = P_r \times \frac{d^n}{G_t G_r} \left(\frac{4\pi}{\lambda} \right)^2 M_l N_f = \bar{\gamma}_b R_b \times \frac{d^n}{G_t G_r} \left(\frac{4\pi}{\lambda} \right)^2 M_l N_f N_0 = \bar{E}_b R_b \times \frac{d^n}{G_t G_r} \left(\frac{4\pi}{\lambda} \right)^2 M_l N_f \quad (2)$$

其中， P_r 为接收端信号功率； $L_{loss} = d^n (4\pi/\lambda)^2$ 表示路径损耗模型； \bar{E}_b 是接收端在给定误比特率下所需要的比特能量， $\bar{\gamma}_b$ 为平均比特信噪比； R_b 为比特传输速率；其它参数详见文献[7]。发送端功率放大器的能耗可以近似为

$$P_{\text{PA}} = (1 + \alpha)P_{\text{out}} = \frac{\xi}{\eta} P_{\text{out}} \quad (3)$$

其中, $\alpha = \xi/\eta - 1$; η 是射频功率放大器的效率因子; ξ 为峰均比(Peak to Average Ratio, PAR); 在 MQAM 调制系统中, $\xi = 3(\sqrt{M} - 1)/(\sqrt{M} + 1)$, 其中, M 是调制电平数。

电路功耗分为发送电路功耗和接收电路功耗, 分别用 P_{tc} 和 P_{rc} 表示。无线通信过程中整个电路能耗可以由下式给出, 其中各参数详见文献[7]。

$$P_{\text{c}} = P_{\text{tc}} + P_{\text{rc}} \approx M_{\text{t}}(P_{\text{DAC}} + P_{\text{mix}} + P_{\text{filt}} + P_{\text{syn}}) + M_{\text{r}}(P_{\text{LNA}} + P_{\text{mix}} + P_{\text{IFA}} + P_{\text{filr}} + P_{\text{ADC}} + P_{\text{syn}}) \quad (4)$$

3 协作MIMO系统能耗分析

我们考虑使用 Alamouti 空时分组码方式实现协作 MIMO 系统的信号发送和接收, 调制方式为 MQAM 调制。在已知传输距离和调制方式的前提下, 求取功放功耗的关键在于获得为满足接收端 BER 要求所需的平均比特能量 \bar{E}_b 。对具有 M_{t} 个发送天线、 M_{r} 个接收天线的协作 MIMO 系统而言, 接收端的瞬时信噪比可表示为^[7]

$$\gamma_b = \frac{\|H\|_F^2}{M_{\text{t}}} \frac{\bar{E}_b}{N_0} \quad (5)$$

其中, \bar{E}_b 表示接收端平均比特能量, $N_0/2$ 为加性高斯白噪声的双边功率谱密度, $\|H\|_F^2$ 为系统信道矩阵 \mathbf{H} 的 Frobenius-2 范数。

$$\|H\|_F^2 = \sum_{i=1}^{M_{\text{t}}} \sum_{j=1}^{M_{\text{r}}} |H_{j,i}|^2 = \sum_{i=1}^{M_{\text{t}}} |h_i|^2 = \sum_{i=1}^{M_{\text{t}}} \{\text{Re}[h_i]\}^2 + \{\text{Im}[h_i]\}^2 \quad (6)$$

其中, 元素 $H_{j,i}$ 表示从发送天线 i 到接收天线 j 的信道复值衰落系数。在信道增益呈现不同的分布衰落时, $H_{j,i}$ 的统计特性也不相同。此处, 我们假设矩阵 \mathbf{H} 中的每项 h_i 均为 ZMCSCG, 方差为 1。式(6)左边可进一步分解为 $2M_{\text{t}}M_{\text{r}}$ 个独立同分布的实高斯分布随机变量的平方和, 其中每个实高斯分布随机变量 ($\text{Re}[h_i]$ 、 $\text{Im}[h_i]$) 的均值为零, 方差为 0.5。显然, $2\|H\|_F^2$ 服从自由度为 $2M_{\text{t}}M_{\text{r}}$ 的中心 $\chi^2(2M_{\text{t}}M_{\text{r}})$ 分布, 其概率密度函数可表示为

$$p_{2\|H\|_F^2}(x) = \begin{cases} \frac{1}{2^{M_{\text{t}}M_{\text{r}}} (M_{\text{t}}M_{\text{r}} - 1)!} x^{M_{\text{t}}M_{\text{r}}-1} e^{-x/2}, & x \geq 0 \\ 0, & x < 0 \end{cases} \quad (7)$$

在 AWGN 条件下, MQAM 的误比特率为^[8]

$$P_b \approx \frac{4}{b} \left(1 - \frac{1}{2^{b/2}}\right) Q \left[\sqrt{\frac{3b\gamma_b}{(M-1)}} \right] \quad (8)$$

其中, $M = 2^b$, $b \geq 2$ 。对 \mathbf{H} 取均值运算, 可得衰落环境中平均误差错率

$$\bar{P}_b = E_{\mathbf{H}} \{P_b\} = \quad (9)$$

根据经典积分公式(10)将式(9)化简,

$$\int_0^{+\infty} x^{2n-1} e^{-\frac{x^2}{2}} Q\left(\frac{x}{\beta}\right) dx = \frac{(n-1)!}{2} \left[1 - (\beta^2 + 1)^{-\frac{1}{2}}\right]^n \sum_{k=0}^{n-1} 2^{-k} \binom{n-1+k}{k} \left[1 + (\beta^2 + 1)^{-\frac{1}{2}}\right]^k \quad (10)$$

可得瑞利信道条件下协作 MIMO 系统的误比特率为

$$\bar{p}_b = \frac{4}{b} \left(1 - \frac{1}{2^{b/2}} \right) \frac{1}{2^{M_t M_r}} \left[1 - (\beta^2 + 1)^{-\frac{1}{2}} \right]^{M_t M_r} \times \sum_{k=0}^{M_t M_r - 1} 2^{-k} \binom{M_t M_r - 1 + k}{k} \left(1 + (\beta^2 + 1)^{-\frac{1}{2}} \right)^k \quad (11)$$

这里, $\beta^2 = \frac{2(M-1)M_t N_0}{3b\bar{E}_b}$ 。当 b 为奇数时, 去掉式(11)中的 $\left(1 - \frac{1}{2^{b/2}} \right)$ 项。

由此, 给定误比特率、 M_t 和 M_r , 根据式(5)、(8)、(11), 即可求取一定误比特率下接收端所需的每比特传输能量 \bar{E}_b 。相应地, 根据式(1)、式(3)和式(4)可得, 每比特传输能量为

$$E_{bt} = (1 + \alpha) \bar{E}_b \times \frac{d^n}{G_t G_r} \left(\frac{4\pi}{\lambda} \right)^2 M_t N_f + P_c / b R_s \quad (12)$$

其中, $bR_s = L/T_{on}$, L 表示传输的总比特数, T_{on} 表示传输持续的时间。在一定的带宽 B 下, 传输时延满足^[7]

$$T_{on} = L/bR_s = LB/b \quad (13)$$

由式(1)可知, 发送一个比特所需的总能耗与传输速率有关。如果符号速率一定, 则每比特的电路能耗随着调制参数 b 的增加而线性下降, 即电路能耗则是 b 的减函数。在一定发射功率下, 多进制调制会导致误比特率的上升; 因而, 要保持一定的误比特率, 势必要提高发射功率, 即功放能耗是调制参数 b 的增函数。对一个固定的距离 d 而言, 一定存在一个最优的调制方式, 使总能耗最小。我们假设信道是瑞利衰落信道且经历大尺度衰落。根据式(11) 中误比特率计算方法求取 \bar{E}_b , 可对不同距离 d 下的可变速率的 SISO 系统、协作 MISO 系统和协作 MIMO 系统的总能耗进行仿真, 寻找出各通信结构下的最优调制参数 b 。

图 1 给出 SISO 系统中调制参数 b 和总能耗之间的关系。由图可见, 在不同传输距离 d 下, SISO 系统的总能耗各异, 且距离 d 越大, 传输所消耗的总能量越多; 在同一传输距离 d 下, 采用不同调制方式时消耗的能量也各不相同。以 $d = 20\text{m}$ 为例, 当调制参数 $b = 4$ 时, 每比特总能耗最小, 且与 $b = 1$ (BPSK 调制) 时的总能耗相比, 有明显的降低; 当 $b > 4$ 时, 随着调制参数的增加, 每比特总能耗呈明显上升趋势。相应地, 表 1 给出 SISO 系统在各个传输距离 d 下的最优调制参数 b 。

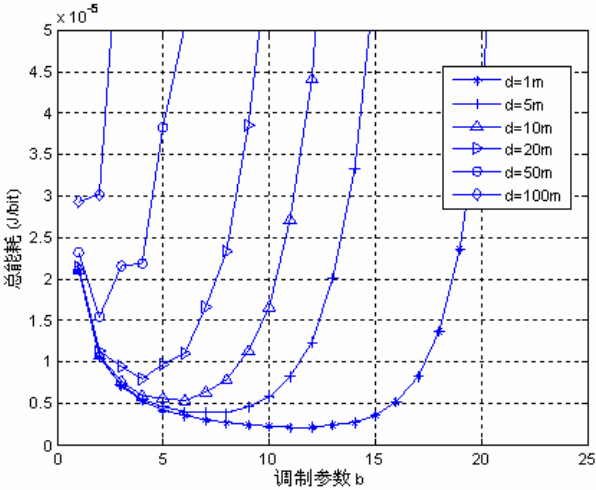


图 1 SISO 系统调制参数和总能耗关系图

表 1 SISO 系统在各个传输距离下的最优调制参数

$d(\text{m})$	1	5	10	20	50	100
最优 b	11	7	6	4	2	1

同样，可得 MISO 系统和 MIMO 系统中调制参数 b 和总能耗的关系，并获得在各个传输距离下的最优调制参数，详见表 2 和表 3。

表 2 MISO 系统在各个传输距离下的最优调制参数

$d(\text{m})$	1	5	10	20	50	100
最优 b	14	10	8	6	4	4

表 3 MIMO 系统在各个传输距离下的最优调制参数

$d(\text{m})$	1	5	10	20	50	100
最优 b	16	12	10	8	6	5

4 自适应调制的协作通信系统的能耗与时延折中

考虑一个分簇的网络，节点间协作通信模型如图 2 所示。该模型中信源与协作节点之间位于不同的簇，它们之间的信息交互不再属于本地通信，而应属远程通信。簇内所有节点之间的距离都按照簇内节点之间的最大距离 d_m 来计算。信源和目的端之间的距离设为 d ，其间 $K-2$ 个簇参与协作通信，簇间距 $d_{ij} \gg d_m$ [14-15]。

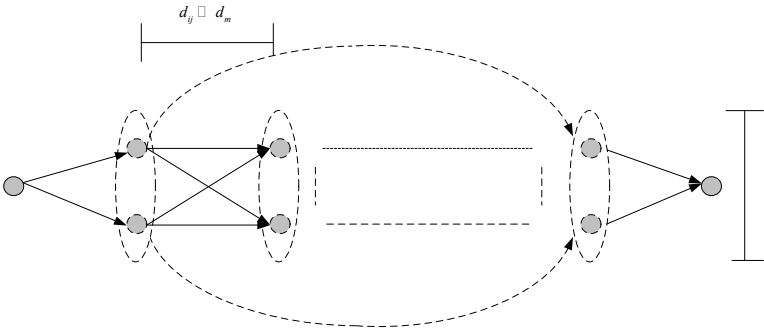


图 2 节点间协作通信模型

4.1 节点间协作通信的能耗与时延分析

为使整个模型尽量简单，我们假设一个簇内参与通信的节点数为 $M = 2$ 。信源与簇 2 间的通信(即为信源的广播过程)，可抽象成 SIMO 模型，其中该模型的功放能耗分析可等同于 SISO 模型；参与协作的相邻簇之间的通信可抽象成协作 MIMO 模型；簇 K 与目的端之间通信可抽象成协作 MISO 模型。

假设簇 i 和簇 j 之间通信距离为 d_{ij} ，调制参数为 b_{ij} ， E_{ij} 表示它们之间的每比特能耗， T_{ij} 表示它们之间传输时延。信源 s 待发送的比特数为 L_1 ，协作簇发送的比特数为 L_i ，

$i = 2, \dots, K-1$ ，且 $L_i = L_{i-1} + L_i^{\text{new}}$ ，其中， L_{i-1} 为簇 i 需要为簇 $i-1$ 转发的比特数， L_i^{new} 表示簇 i 自己新产生的待传输的比特数。网络总能耗表示为

$$E_{\text{sum}} = \sum_{i=1}^{K-1} L_{ij} E_{ij}, \quad j = i + 1 \quad (14)$$

网络总时延表示为

$$T_{\text{sum}} = \sum_{i=1}^{K-1} T_{ij}, \quad j = i + 1 \quad (15)$$

4.2 自适应调制的协作通信网络的性能优化

对于一个自适应调制的协作通信系统，系统的能耗与时延性能都与调制参数 b 有着密不可分的关系。对于不同的传输距离，基于每比特总能耗最小的原则选择最优的调制参数 b 。在实际通信过程中，系统性能的优化往往受到多种条件的制约，比如传输时延、总能耗、发送功率等等，甚至还要受到通信协议和网络拓扑等的影响。基于此，本文我们提出发送功率受限的协作通信系统性能优化策略（ δ 是权值因子），即

$$\begin{aligned} & \text{Minimize } T_{\text{sum}} + \delta E_{\text{sum}} \\ & \text{s.t. } P_{\text{out}} \leq P_{\text{max}} \\ & 0 \leq L_i^{\text{new}} \leq L_{\text{max}}, i = 2, \dots, K-1 \end{aligned} \quad (16)$$

仿真中，假设网络节点分布如图 2 所示，信源与目的节点之间距离为 $d=1000\text{m}$ ，它们之间相隔 9 个节点簇，即 $K=11$ 。所有簇之间等距，簇间距 $d_{ij} = d/(K-1) = 100\text{m}$ 。每个节点在 BPSK 调制下的最大传输距离为 180m ，根据链路预算公式，可得最大发送功率为 222.5mW 。如图 3 所示的系统结构示意图中，簇 2、簇 5 和簇 6 参与了协作。随着参与协作的簇的不同，协作簇之间距离也各异。此时，网络将呈现不同的拓扑结构。

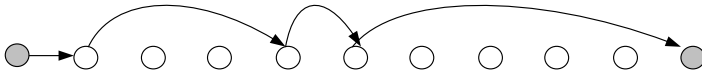


图 3 部分簇参与协作的系统结构图

在最大发送功率限制下，各种通信结构下的最大传输距离与调制参数 b 之间受到相互制约，信源只能与簇 2 进行一跳通信，协作簇与目的端之间的通信(2×1 MISO 结构)的通信距离不能超过 5 个簇间距，两个协作簇之间的通信距离不能超过 12 个簇间距。因而，MIMO 簇间协作最多占用 8 个簇间距，最少占用 4 个簇间距。使用穷举法，搜索最优的网络拓扑结构及其最优的能量与时延关系，图 4 给出不同拓扑下的总能耗与总时延性能图，其中参数 $\delta=1$ ，X 轴表示 MIMO 结构分段数。

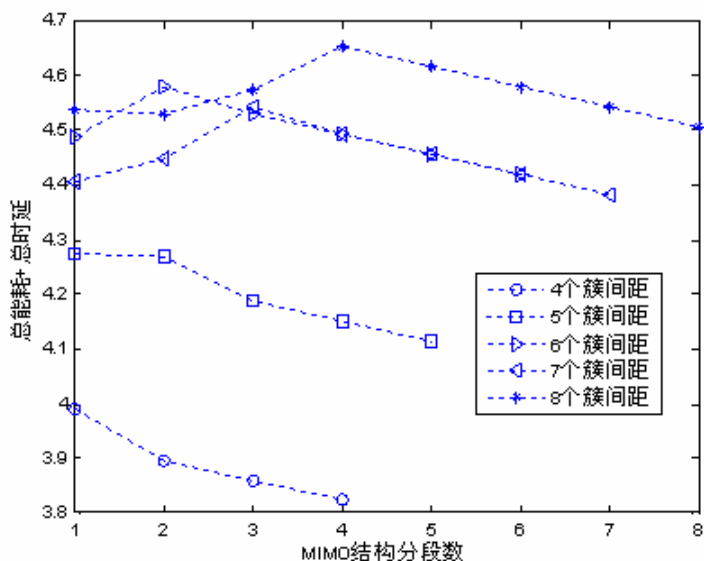


图4 不同拓扑下的协作网络性能

从图4中可见,当协作MIMO通信横跨4个簇间距时,系统性能相较于其他情况更优。其中,当分段数为4,即如图5的协作结构时性能最优。

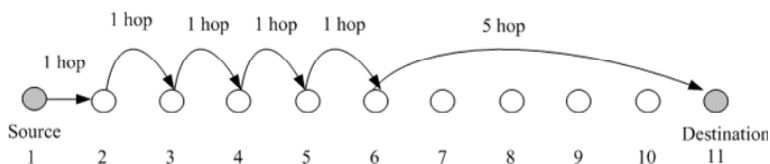


图5 协作网络最优拓扑结构

5 结论

协作分集是空间分集的一种新形式,它利用无线信道的广播特性,选择部分侦听到信源发送信息的节点作为中继,辅助信源进行数据包的传送,本文分析结果表明,协作分集无线通信与无协作分集的无线通信相比,能够明显提高网络的能量有效性,降低网络功耗。此外,MAC接入方式对网络能耗也有重要影响,作者将基于跨层的方法对随机接入方式下的协作分集无线通信能耗作进一步的分析 and 研究。

参 考 文 献

- [1] A. Sendonaris, E. Erkip, B. Aazhang. User cooperation diversity-part I and part II [J]. IEEE Transactions on Communication, 2003, 51(11): 1927~1948.
- [2] V. Stankovic, A. Host-Madsen, Xiong Zixiang. Cooperative diversity for wireless ad hoc networks [J]. IEEE Signal Processing Magazine, 2006, 23(5): 37~49.

- [3] A. Scaglione, D.L. Goeckel, J.N. Laneman. Cooperative communications in mobile ad hoc networks [J] . IEEE Signal Processing Magazine, 2006, 23(5): 18~29.
- [4] M. Dianati, Ling Xinhua, K. Naik, Shen Xuemin. A node-cooperative ARQ scheme for wireless ad hoc networks[J].IEEE Transactions on Vehicular Technology, 2006, 55(3): 1032~1044.
- [5] JI Wei, ZHENG Baoyu. Enhanced Cooperative Packet Retransmission Employing Joint Cooperative Diversity and Symbol Mapping Diversity. Journal of Zhejiang University -Science A. 2008, 9(8):1090-1098.
- [6] Hong Yao-Win, A. Scaglione. Energy-efficient broadcasting with cooperative transmissions in wireless sensor networks [J]. IEEE Transactions on Wireless Communications, 2006, 5(10): 2844~2855.
- [7] Cui Shuguang, A.J. Goldsmith, A. Bahai. Energy-efficiency of MIMO and Cooperative MIMO techniques in sensor networks [J].IEEE Journal on Selected Areas in Communications, 2004, 22(6): 1089~1098.
- [8] Cui Shuguang, A.J. Goldsmith. Cross-layer design of energy-constrained networks using cooperative MIMO techniques[J]. Signal Processing, 2006,86(8):1804~1814.
- [9] Cui S., A.J. Goldsmith. Cross-layer optimization of sensor networks based on cooperative MIMO techniques with rate adaptation. 2005 IEEE 6th Workshop on Signal Processing Advances in Wireless Communications, 5-8 June, 2005, Pages: 960~964.
- [10] S.K. Jayaweera. Energy efficient virtual MIMO-based cooperative communications for wireless sensor networks . Proceedings of 2005 International Conference on Intelligent Sensing and Information Processing, 4~7 Jan, 2005, Pages: 1~6.
- [11] S.K. Jayaweera, M.L. Chebolu. Virtual MIMO and distributed signal processing for sensor networks-an integrated approach . 2005 IEEE International Conference on Communication (ICC 2005), 16~20 May, 2005, 2: 1214~1218.
- [12] S.K. Jayaweera. Virtual MIMO-based cooperative communication for energy - constrained wireless sensor networks [J]. IEEE Transactions on Wireless Communications, 2006, 5(5): 984~989.
- [13] 陈贞贞, 杨晨阳. 无线传感网物理层协作分集的性能分析. 电子与信息学报. 2007, 29(11): 2569-2573.
- [14] Cui, S., Goldsmith, A. J., Bahai, A. Energy-Constrained Modulation Optimization. IEEE Transactions on Wireless Communications. 2005, 4(5):2349-2360.
- [15] Cui, S., Goldsmith, A. J. Cross-Layer Design of Energy-Constrained Networks using Cooperative MIMO Techniques. Signal Processing. 2006, 86:1804-1814.

作者简介:

季薇, 女, 1979 年生, 江苏淮安人, 2008 年 12 月毕业于上海交通大学电子工程系获博士学位, 现为南京邮电大学通信与信息工程学院讲师, 主要研究方向为无线网络中的协作分集与协作通信。

郑宝玉, 男, 1945 年生, 福建省闽侯人, 南京邮电大学教授、博士生导师, 上海交通大学兼职教授、博士生导师, 中国通信学会通信理论与信号处理专业委员会主任委员, 主要研究方向是智能信号处理、通信信号处理和量子信号处理等。

江若宜, 女, 1986 年生, 福建省人, 南京邮电大学通信与信息学院硕士, 主要研究方向为无线协作通信中的能量有效性问题。

时间选择性衰落信道的基扩展模型比较

陈东华 蔡灿辉

(华侨大学信息科学与工程学院 福建 泉州 362021)

摘要: 时间选择性信道时域响应的直接估计通常需要发端发送大量的训练序列或导频; 采用基扩展模型只需用少量参数即可表示时变信道, 降低了时变信道估计的复杂度和导频率, 本文研究了三种常用的基扩展模型: 指数基、多项式基、KL 基扩展模型, 比较研究了三种基扩展模型的逼近精度, 同时研究了基数目对模型精度的影响, 数值结果表明, KL 基模型的逼近精度最高, 对指数基过采样可提高其逼近精度。

关键词: 信道估计; 时变信道; 基扩展模型;

Comparative Study on Basis Expansion Models in Time-Selective Fading Channels

Chen Dong hua Cai Can hui

(College of Info. Science and Engineering, Huaqiao University, Quanzhou, 362021, china)

Abstract: Direct estimation of time-selective channels in time domain involves large training overhead and complexity. Time-selective channels' representation using only a little number of BEM coefficients can reduce the complexity and overhead in channel estimation. CE-BEM, P-BEM and KL-BEM are three important BEM, the approximate precisions of which are studied comparatively. The impactions of the number of BEM coefficients on approximate precisions are also studied. Numerical results show that KL-BEM is the most precise BEM and the OCE-BEM is superior on CE-BEM in approximate precision.

Keywords: Channel estimation, time varying channels, basis expansion model.

1 引言

未来的无线通信系统要求能在高速移动的环境下提供高速的数据传输。数据率的提高要求有更宽的带宽, 而带宽的增加会使采样间隔小于信道的时延扩展, 这样就产生了多径环境下的频率选择性衰落。同时移动台和基站之间的高速相对移动带来多普勒效应, 多普勒效应使得传输信道随时间而快速变化, 也就是信道的时选择性衰落。发射机和接收机之间的相对运动和频偏都会导致信道快速时变, 而且随着载频和移动速度的提高, 信道的时变特性越来越明显。未来高移动性的宽带通信系统将会受到时-频双选择性衰落的严重影响, 因此近些年来双选信道的估计越来越受到关注^[1, 6], 双选择性信道估计的精度对高速宽带移动通信系统的性能至关重要。

时间选择性信道的直接估计需要估计大量的信道抽头样值, 因为每一个抽样时刻都要估计, 这一方面要求发射机要占用大量的带宽传输训练序列或导频序列, 另一方面使得接收机的信道估计复杂度增加。因此寻求时变信道的简化表示变得十分重要, 最近, 利用基扩展模型 (BEM) 来近似无线信道的时变特性得到广泛的关注, 基扩展模型用有限个数的不变基系数表示传输块内信道的时变, 文献[2]采用复指数基扩展模型对时变信道进行近似, 文献[3]基于泰勒级数展开用少量多项式函数来近似时变信道, 文献[4]采用 KL 正交扩展实现相同的目标。上述三种最常见的时变信道基扩展模型逼近的性能比较, 对实际应用中的指导作用具有十分重要的意义, 本文从三种基扩展模型的逼近精度, 基数目对模型精度的影响等方面对三种模型进行了比较研究。

2 有限参数信道模型传输系统^[3,7]

假设无线信道最大多径时延扩展为 L , 系统输入输出离散时间等效基带信号可表示为

$$r(n) = \sum_{l=0}^{L-1} h(n,l)s(n-l) \quad (1)$$

其中, $h(n,l)$ 为 n 时刻第 l 条路径的时域信道响应, $r(n)$ 和 $s(n)$ 分别为 n 时刻的收发信号; 为了克服频率选择性引起的码间串扰, 通常采用块传输系统, 块传输系统如 OFDM, 单载波频域均衡系统, 在每一个传输块前增加保护间隔如循环前缀 (CP), 只要保护间隔长度大于信道时延扩展 L , 即可消除块间干扰。假设每一传输块的长度为 N , 循环前缀长度 $G \geq L$, 将块内的 N 个接收样值和发送样值分别表示为矢量形式 $\mathbf{r} = [r(0) \cdots r(N-1)]$, $\mathbf{s} = [s(0) \cdots s(N-1)]$, 则除去循环前缀后接收信号的等效矢量表示为

$$\mathbf{r} = \mathbf{H}\mathbf{s} \quad (2)$$

其中, \mathbf{H} 为 $N \times N$ 时域信道矩阵。如果直接估计时域信道矩阵 \mathbf{H} , 则需要估计 $N \times L$ 个信道样值, 显然所需要大量训练序列和较高的计算复杂度。为降低复杂度, 通常采用基扩展表示的有限参数模型来近似是时变信道, 下面介绍三种典型的时变信道的基扩展模型。

3 基扩展模型^[3,7]

时变信道的有限参数基扩展模型如下

$$h(n,l) = \sum_{m=0}^{M-1} c(m;l)u_m(n) \quad n \in [0, N-1] \quad (3)$$

其中 $c(m;l)$ 表示长为 N 的传输块时间内第 l 条路径的第 m 个有限参数模型系数, $u_m(n)$ 是对应基函数在 n 时刻的取值, 将 (3) 式表示为矢量形式

$$\mathbf{h}_l = \mathbf{U}_l \mathbf{c}_l \quad (4)$$

其中, $\mathbf{h}_l = [h(0,l) \cdots h(N-1,l)]^T$, $\mathbf{c}_l = [c(0;l) \cdots c(M-1;l)]^T$, \mathbf{U}_l 表示 $N \times M$ 的基函数组成的矩阵, 选择不同的基函数就得到不同的基扩展模型。

公式 (3) 用 M 个基来捕获每一径上的时变特性, 也就是说每一径的时变信道, 仅用 M 个基的系数就可以表示, 通常 M 的取值很小, 那么用较少数量的 $M \times L$ 个系数 \mathbf{c} 就可以描述整个块内时变多径信道, 而一个块的长度 N 通常很大, 从几百到几千, 这样估计的参数数量大大降

低，信道估计的复杂性也大大降低。

3.1 复指数基扩展模型^[2,7]

复指数基即傅立叶基， $h(n,l)=\sum_{m=0}^{M-1}c(m;l)\exp(j2\pi(m-M/2)n/N)$ ，基个数 $M=2\lceil f_{\max}NT_s \rceil+1$ ， T_s 为采样间隔， f_{\max} 为最大多普勒频移。用 BEM 基扩展模型来表示时变信道，相当于在多普勒扩展频谱上进行频域采样，用这些采样频点的值来近似表示一定长度数据块内信道的变化，当块长等于 N 时，多普勒频谱采样间隔（即多普勒频率分辨率） $\nu_1=1/N$ ，这里采用等间隔采样，第 m 个采样频率 $\nu_m=(m-M/2)/N$ ，但是当信道变化较慢时，多普勒扩展带宽很窄， M 个实际频域采样点将超出多普勒扩展带宽，部分基函数将不能真实表示多普勒扩展频谱的频域采样，造成“多普勒泄漏效应”，为克服指数基扩展模型的缺点，文献[5]提出过采样 BEM（OCE-BEM），OCE-BEM 周期取为块长 N 的 K 倍，即 $\nu_1=1/(N\times K)$ ，当 $K=1/(f_{\max}NT_s)$ 时，采样点的位置正好在 doppler 扩展带宽的持续范围内，如图 1 所示。从图 1 中可以看到，当 $f_{\max}NT_s<1$ 时，标准的 BEM 基频率的采样位置与实际信道变化的频率范围存在偏差，而过采样 BEM 基频率的位置则在实际信道变化的频率范围内，能够更好地近似实际信道的变化。

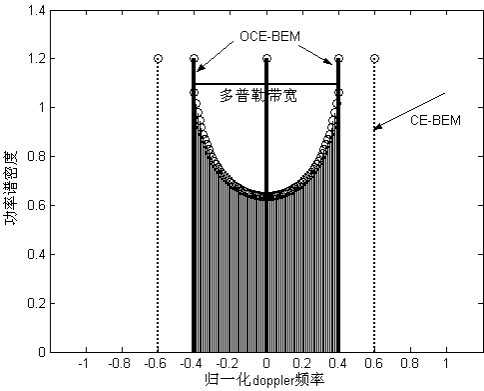


图 1 CE-BEM 的多普勒泄漏示意

3.2 多项式基扩展模型^[3]

多项式基扩展模型能克服 CE-BEM 的多普勒泄漏效应，多项式基扩展模型由泰勒级数展开得到，任意时刻 n 时的信道脉冲响应 $h(n,l)$ 可近似表示为 $h(n,l)=\sum_{m=0}^{M-1}c(m;l)n^m$ 。

3.3 KL基扩展模型^[4,5]

上述两种基扩展模型都没有任何意义下的最优性能，KL 扩展则是在最小均方误差准则下的最优有限参数时变信道模型。令 \mathbf{R}_l 表示第 l 径的时域信道相关矩阵， $\mathbf{R}_l=E[\mathbf{h}_l\mathbf{h}_l^H]$ ，对 \mathbf{R}_l 进行奇异值分解， $\mathbf{R}_l=\boldsymbol{\Psi}_l\boldsymbol{\Lambda}_l\boldsymbol{\Psi}_l^H$ ，其中 $\boldsymbol{\Lambda}_l=diag(\{\lambda_{m,l}\}_{m=0}^{N-1})$ 为 \mathbf{R}_l 的按从大到小排序的特征值，设当 $m>M$ 时， $\lambda_{m,l}\approx 0$ ， $\boldsymbol{\Psi}_l$ 为特征值对应的特征矢量组成的矩阵， $\mathbf{U}_l=\boldsymbol{\Psi}_l^{(:,1:M)}$ ，即 \mathbf{U}_l 取自 $\boldsymbol{\Psi}_l$

的前 M 列时, 因此基矩阵仅依赖于信道自相关矩阵 \mathbf{R}_l 。每径信道抽头 $\{h_l(n)\}_{l=0}^{L-1}$ 为具有 Jakes Doppler 谱的相互独立、零均值复高斯分布的随机过程, 因此第 l 径信道抽头的信道自相关为 $E[h_l(n)h_l^*(m)] = J_0(2\pi f_{\max}(n-m)T_s)$, 其中 $J_0(\cdot)$ 为第一类零阶贝塞尔函数, 当系统传输速率确定后, 信道自相关仅依赖于移动台和基站的相对移动速度。

4 模型逼近性能比较

由式 (1), 基扩展函数的系数的最小二乘解 $\mathbf{c}_l = \mathbf{U}(\mathbf{s}_l\mathbf{U})\mathbf{r}$, 其中, \mathbf{s}_l 为 \mathbf{s} 的循环 l 位移矢量。

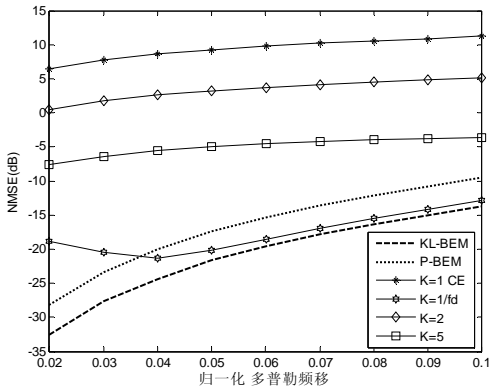


图 2 模型阶 $M=3$ 时三种有限参数模型的逼近精度曲线

4.1 三种基扩展模型的逼近精度比较

本小节, 我们对双选信道的三种模型进行数值仿真比较, 采用第二节介绍的块传输系统, $N=64$, $G=16$, 图 2 为三种有限参数模型的逼近精度比较, 横轴为归一化 doppler 频移 $f_d = f_{\max}(N+G)T_s$; 纵轴为信道估计归一化均方误差 NMSE, 定义为 $\text{NMSE}(n) = \sum_l E\{|h(n,l) - \hat{h}(n,l)|^2\} / \sum_l E\{|h(n,l)|^2\}$; 固定模型阶数为 $M=3$, $L=8$, 由图可见, 三种模型中, KL 模型的近似误差最小, 即逼近精度最高, 而过采样因子为 $K=1/(f_{\max}NT_s)$ 的 OCE-BEM 其次, 没有过采样的 CE-BEM 最差, 但过采样的 CE-BEM 基函数不再保持正交, 图中同时给出了过采样因子 $K=2, 5$ 的情况。

4.2 模型阶对逼近精度的影响

图 3 为不同模型阶下 P-BEM 和 KL-BEM 模型的逼近精度, 由图可见, 随着模型阶的增加, 各种基扩展模型的逼近性能也随之增加, 而且在相同的模型阶下, KL-BEM 性能最优, P-BEM 差于 KL-BEM。

从三种模型的逼近机理我们分析一下上述原因, KL 基是利用信道的二阶统计特性使得 MSE 最小意义下的一种有限参数信道模型, 它可以根据信道实际的归一化多普勒来构造其基函数, 而多项式模型使用泰勒级数展开近似得到的, 因此逼近的 MSE 不能达到最优, 复指数基扩展模型相当于在多普勒扩展频谱上进行频域采样, 用这些采样频点的值来近似表示一定长

度数据块内信道的变化,因此也不是一种最优逼近。过采样因子 $K=1/(f_{\max}NT_s)$ 的 OCE-BEM 之所以比基本 CE-BEM (未过采样) 近似精度高,主要是由于当信道变化比较慢时, $f_{\max}NT_s$ 较小,多普勒频谱带宽很窄,多普勒扩展频谱上进行频域采样的范围超过实际的多普勒频谱带宽,部分采样值不在实际的多普勒频谱带宽范围内,因此与真实信道差别较大,而过采样 CE-BEM 基频率的位置则在实际信道变化的频率范围内,能够更好地近似实际信道的变化。

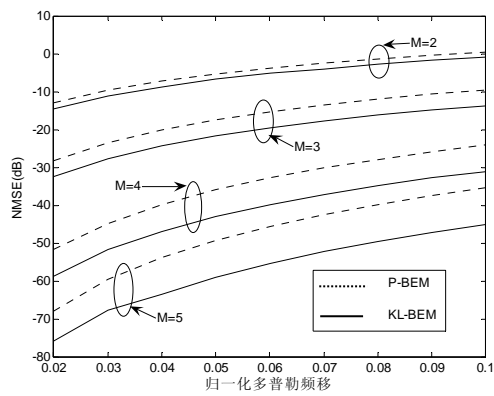


图 3 不同模型阶下 P-BEM 和 KL-BEM 模型的逼近精度

5 结束语

本文系统分析了双选择性信道的有限参数基扩展模型表示,双选信道通过有限参数模型表示,可以有效减少收发信机的信道估计算法,文中我们仿真比较研究了 CE-BEM, OCE-BEM, P_BEM, KL_BEM 几种最常用的时变信道基扩展模型的逼近精度,同时仿真比较了基数目对模型精度的影响,数值结果表明,KL 扩展基模型的逼近精度最高,对 CE-BEM 模型进行分数倍过采样可提高复指数基扩展模型的精度。

参 考 文 献

- [1] Schniter P. Low complexity estimation of doubly selective channels [C]. Proc of IEEE Workshop Signal Processing Advances in Wireless Commun. Rome, June, 2003:200-204.
- [2] Ma Xiaoli, Giannakis G B, Ohno S. Optimal training for block transmissions over doubly selective wireless fading channels [J]. IEEE trans on Signal Processing, 2003, 51 (5): 1351-1366.
- [3] Deva K. Borah, and Brian D. Hart. Frequency-Selective Fading Channel Estimation with a Polynomial Time-Varying Channel Model [J]. IEEE trans on Commun,1999,47(6):862-873.
- [4] M. Visintin, Karhunen-Loeve expansion of a fast Rayleigh fading process [J]. Electron. Lett.,1996, 32(18):1712-1713.
- [5] Qen Qin and Qicong Peng. Joint channel and phase noise estimation in ofdm using KL expansion [J]. International Journal of Electronics and Communications, 2007
- [6] 刘英男等. 基于分数倍基扩展模型的双选信道估计方法 [J]. 北京大学学报(自然科学版),2008, 44(1):87-92
- [7] Giannakis G B, and Tepedelenioglu C. Basis expansion models and diversity techniques for blind identification and equalization of time-varying channel [J]. Proc IEEE,1998,86:1969-1986.

多小区OFDMA系统下行链路的分布式子载波分配

邱 涛 贺志强 牛 凯 田宝玉

(北京邮电大学 信息与通信工程学院, 北京 100876)

摘 要: 本文研究了采用正交频分多址(OFDMA)系统下行链路的分布式子载波分配问题, 综合考虑了物理层上的约束条件(用户的误比特率性能)和媒体接入层的约束条件(用户的最小速率需求), 提出了一种两级分布式子载波分配方案, 就是将子载波分配分成小区间子载波分配和小区内子载波分配两步, 小区间子载波分配通过相邻小区的信息传递可以有效的降低OFDMA系统的主要干扰——小区间的同频干扰, 在小区间子载波分配的基础上, 小区内子载波分配充分考虑用户的业务需求, 根据用户的优先级分配子载波。仿真结果表明, 与软频率复用方案相比较, 新算法虽然引入了一定的复杂度, 但是可以在满足用户的最小速率需求的同时, 大大提高系统吞吐量。

关键词: 子载波分配; 最小速率需求; OFDMA

Distributed Sub-carrier Allocation Scheme for Multi-cell OFDMA System Downlink

Qiu Tao He Zhi qiang Niu Kai Tian Bao yu

(School of Information and Communication Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, China)

Abstract: The sub-carrier allocation problem for the downlink of multi-cell orthogonal frequency division multiple access system is studied. A two-level distributed sub-carrier allocation algorithm is proposed, it jointly takes the constraints posed by the physical layer (in terms of bit error rate (BER)) and the constraints posed by the media access control layer (in terms of minimum rate requirement of user). The proposed algorithm includes two procedures, that is, inter-cell subcarrier allocation and intra-cell subcarrier allocation. The inter-cell sub-carrier allocation can efficiently reduce the inter-cell interference which is the main interference in OFDMA systems through passing messages between neighbor cells. Based on the inter-cell sub-carrier allocation, the intra-cell sub-carrier allocation can further reduce the number of users that don't satisfied the minimum rate requirement and greatly increase system throughput. It is then compared with the conventional soft frequency reuse schemes. Simulation results show that the proposed algorithm can achieve high throughput and performs much better than soft frequency reuse schemes in terms of the number of users that unsatisfied the minimum rate requirement.

Keywords: sub-carrier allocation; minimum rate requirement; OFDMA

1 绪论

正交频分复用（OFDM）作为一种多载波调制技术，具有频谱利用率高，抗频率选择性衰落或窄带干扰等优点，已成为下一代移动通信技术的热点。由于无线信道的时变性和衰落特性，OFDM 系统中各个子载波条件各不相同。一个子载波若被某些用户占用，会比其它用户占用带来更大的收益。

在 OFDMA 系统中，子载波分配是不可或缺的一个重要环节，子载波分配就是要根据信道状况进行合理的分配，有效的利用频率资源，减小同频信道的干扰。如何更好的提高网络中的频谱利用率，如何更好的满足用户之间的公平性，都需要在子载波分配中进行重点考虑。

对于单小区 OFDMA 系统，文献[1]给出了“将每个子载波分配给信道增益最好的用户可以使多用户 OFDM 系统的数据速率达到最大”的结论。对于多用户 OFDM 系统，由于不同用户的信号经历了相互独立的信道衰落，因此可以通过动态分配子载波充分利用多用户分集[2]。

近期，许多学者已经提出了很多用于解决多小区下行链路资源分配问题的方法^[3,8]。传统的软频率复用方案(文献[3]~[5])在系统频谱效率和用户公平性之间获得折中，但是由于它们多数是静态或者半静态的频率复用方案，限制了系统吞吐量。文献[6]提出了一种通过获得所有小区全部信息来最大化系统总吞吐量的方案，但是并没有充分考虑用户的 QoS 需求，而且是一种半分布式的分配方案。

本文综合考虑了物理层上的约束条件(用户的误比特率性能)和媒体接入层的约束条件(用户的最小速率需求)，提出了一种两级分布式子载波分配方案，就是将子载波分配分成小区间子载波分配和小区内子载波分配两步，小区间子载波分配通过相邻小区的信息传递可以有效的降低 OFDMA 系统的主要干扰——小区间的同频干扰，在小区间子载波分配的基础上，小区内子载波分配充分考虑用户的业务需求，根据用户的优先级分配子载波。仿真结果表明，与软频率复用方案相比较，新算法虽然引入了一定的复杂度，但是可以在满足用户的最小速率需求的同时，大大提高系统吞吐量。

本文的组织如下：第二节本文首先给出了 OFDMA 系统模型并将系统优化问题分解为两个子问题；在第三节中，给出了相应子问题的启发式解决方案，并在第四节中通过仿真验证了算法的优点；第五节给出了本文的总结。

2 系统模型和问题描述

2.1 系统模型

考虑一个系统中有 M 个小区，系统总用户数为 K ，随机的分布在网络中，系统的可用带宽为 W 。给出以下的假设：

- (1) 在一个小区中，一个子载波只能分配给一个用户，因此不存在小区内干扰；
- (2) 每个小区只受到相邻小区的干扰；
- (3) 相邻小区根据相互间的干扰对子载波进行复用；
- (4) 每个小区在每个子载波上的发射功率是相等的；

(5) 基站可以获得完美的信道状态信息。

令 G_{mk}^n 表示用户 k 和基站 m 在子载波 n 上的信道增益。因此，给定网络功率矢量 \mathbf{P} ，小区 m 中的用户 k 的 SINR 可以表示为

$$\gamma_{mk}^n(\mathbf{P}) = \frac{G_{mk}^n P_m^n}{\sum_{l=1, l \neq m}^M G_{lk}^n P_l^n + \sigma^2} \quad (1)$$

其中， P_m^n 表示小区 m 在子载波 n 上的传输功率， σ^2 是噪声功率。那么可达速率可以表示为

$$R_{mk}^n(\mathbf{P}) = \frac{W}{N} \log_2 \left(1 + \frac{\gamma_{mk}^n(\mathbf{P})}{\Gamma} \right) \quad (2)$$

其中 $\Gamma = -\ln(5\text{BER}_k^n)/1.5$ ， BER_k^n 是误比特率要求 0。定义 $K \times N$ 的子载波分配矩阵 \mathbf{A} ，如果子载波 n 分配给用户 k ，则 $a_k^n = 1$ ，否则， $a_k^n = 0$ 。

2.2 系统优化

从系统优化的角度看，优化目标是在系统功率和用户的最小速率需求的约束下通过动态的子载波分配最大化系统吞吐量。

$$\begin{aligned} & \underset{\mathbf{A}}{\text{maximize}} \sum_{m=1}^M \sum_{k \in U_m} \sum_{n=1}^N a_k^n R_{mk}^n \\ & \text{subject to} \quad \sum_{k \in U_m} a_k^n = 1, \forall n \\ & \quad \sum_{n=1}^N R_{mk}^n \geq R_k^{\min} \end{aligned} \quad (3)$$

其中 U_m 表示小区 m 中的用户集合， R_k^{\min} 表示用户 k 的最小速率需求。

为了获得全局最优解，系统需要知道所有用户的信道状态以及业务信息。因此，基站间需要传递的信息量是非常大的。本文所提的两级分布式方案可以有效的降低系统开销以及计算复杂度。

2.3 小区间子载波分配问题描述

小区间子载波分配的主要目的是降低小区之间的相互干扰。为了更好的阐述小区间子载波分配问题，令 N_m^{\min} 表示小区 m 的最小子载波数需求，定义 $M \times N$ 小区间子载波分配矩阵 \mathbf{B} ，如果子载波 n 分配给小区 m ，则 $b_m^n = 1$ ，否则， $b_m^n = 0$ 。

现在，小区间的子载波分配问题可以描述为：

$$\begin{aligned} & \underset{\mathbf{B}}{\text{maximize}} \sum_{m=1}^M \sum_{n=1}^N b_m^n R_m^n \\ & \text{subject to} \quad N_m > N_m^{\min}, \forall m \\ & \quad b_m^n \in \{0, 1\}, \forall m, n \end{aligned} \quad (4)$$

其中 R_m^n 表示小区 m 在子载波 n 上的可达速率，小区间子载波分配试图寻找使系统吞吐量最大化的分配矩阵 \mathbf{B} 。

2.4 小区内子载波分配问题描述

小区间子载波分配是子载波的初步分配，在此基础上，分配给每个小区的子载波集合是已知的，那么小区内子载波分配的目的就是捕获用户业务分集以及信道的衰落分集。为了更好的阐述小区内子载波分配算法，以小区 m 为例，令 K_m 表示小区 m 的用户集合， S_m 表示分配给小区 m 的子载波集合。定义 $K_m \times N$ 小区内子载波分配矩阵 \mathbf{C} ，如果子载波 n 分配给用户 k ， $c_k^n = 1$ ，否则， $c_k^n = 0$ 。

那么，小区内子载波分配问题可以描述为：

$$\begin{aligned} &\underset{A_m}{\text{maximize}} \quad \sum_{k \in U_m} \sum_{n \in S_m} a_{mk}^n r_k^n \\ &\text{subject to} \quad \sum_{k \in U_m} a_{mk}^n = 1, \forall n \in S_m \\ &\quad \quad \quad R_k \geq R_k^{\min}, \forall k \in U_m \end{aligned} \tag{5}$$

其中 r_k^n 表示用户 k 在子载波 n 上的可达速率， R_k 是用户 k 的可达速率， $R_k = \sum_{n \in S_m} r_k^n$ 。

3 所提算法

小区间子载波分配问题和小区内子载波分配问题都是NP-hard的非线性整数优化问题，而且没有有效地计算方法来获得其最优解 0。因此，本章我们提出两种次优的启发式算法来解决这两个子载波分配问题。所提启发式算法属于“局部搜索”方法，在非线性整数优化问题中是一个非常有效的解决方案 0。

3.1 小区间子载波分配快速算法

所提算法的基本思想是通过相邻小区的信息传递决定各小区可用的子载波集合。相邻小区需要传递的信息包括相邻小区在每个子载波上的 SINR 值。在本文中，小区 m 在子载波 n 上的 SINR 指的是本小区所有用户在子载波 n 上的 SINR 的最小值，比如 $\gamma_m^n = \underset{k \in U_m}{\text{minimize}} \gamma_{mk}^n$ ，其中 γ_m^n 表示小区 m 在子载波 n 上的 SINR。

小区间子载波分配算法通过比较不同小区在每个子载波上的SINR值来决定小区是否放弃使用某个子载波。以小区 m 为例来详细阐述小区间子载波分配算法，如 表所示。

表 1 小区间子载波分配算法

1. 信息搜集： 搜集相邻小区在每个子载波上的 SINR 信息。
2. 初始化： N_m^{\min} = 预先设定的值， S_m 包括所有的子载波。
3. 子载波分配： (1) 对于子载波 n ，对 $[\gamma_{1n}^n, \gamma_{2n}^n, \dots, \gamma_{Mn}^n]$ 进行升序排列，并表示为 $[\gamma_{(1)n}^n, \gamma_{(2)n}^n, \dots, \gamma_{(M)n}^n]$ ； (2) 每个小区根据 $[\gamma_{(1)n}^n, \gamma_{(2)n}^n, \dots, \gamma_{(M)n}^n]$ 依次放弃子载波 n 。如果小区 m 放弃子载波 n 不能满足最少子载波数需求 N_m^{\min} ，跳到序列中的下一个小区；否则计算 Ω_m ，如果 $\Omega_m > 0$ 使 $b_m^n = 0$ 并更新 $S_m = S_m - \{n\}$ 和 $[\gamma_{(1)n}^n, \gamma_{(2)n}^n, \dots, \gamma_{(M)n}^n]$ ，否则使 $b_m^n = 1$ 并跳到下一个子载波； (3) 直到所有子载波分配完毕。

实际上，算法是逐个子载波进行，对于每个子载波，当小区放弃使用时计算 Ω_m ， Ω_m 表示小区 m 对于系统吞吐量提高的贡献，相当于降低了的小区间干扰与子载波低利用率之间的折中。算法执行分为两个循环，即子载波循环和小区循环，所以小区间子载波分配算法的复杂度为 $O(M^2 \times N)$ 。

3.2 小区内子载波分配快速算法

当每个小区接收到分配的子载波集合，就需要按照用户的最小速率需求将子载波与用户进行配对。本算法首先根据用户的几何 SINR 将用户分为性能较好用户和性能较差用户，然后按照先性能较差用户后性能较好用户的顺序分配子载波。还是以小区 m 为例来详细阐述小区内子载波分配算法，如表 2 所示。

表 2 小区内子载波分配算法

1. 初始化： R_k^{\min} = 预先设定的值， $\forall k \in U_m$ 。假设所有子载波都未分配并对每个用户的 $[\gamma_{mk}^1, \gamma_{mk}^2, \dots, \gamma_{mk}^N]$ 进行降序排列，表示为 $[\gamma_{mk}^{(1)}, \gamma_{mk}^{(2)}, \dots, \gamma_{mk}^{(N)}]$ ；
2. 性能较差用户子载波分配： (1) 对小区 m 中性能较差用户计算 $R_k^{\min} - R_k$ ； (2) 找出 $R_k^{\min} - R_k$ 最大的用户 k^* ，根据 $[\gamma_{mk^*}^{(1)}, \gamma_{mk^*}^{(2)}, \dots, \gamma_{mk^*}^{(N)}]$ 将子载波分配给用户 k^* ，如果子载波被其他用户占用，跳到序列中下一个子载波。更新 $R_{k^*}^{\min} - R_{k^*}$ ，跳到(2)直到满足所有用户的最小速率需求或者所有子载波分配完毕；
3. 性能较好用户子载波分配： 与性能较差用户相同；
4. 剩余子载波分配 如果还有子载波未分配，将每个子载波分配给 SINR 值最大的那个用户，直到所有子载波分配完毕。

实际上，小区内子载波分配根据用户的优先级进行子载波分配，性能较差用户的优先级高于性能较好用户，而且 $R_k^{\min} - R_k$ 越大，优先级越高。可以证明，小区内子载波分配算法的复杂度为 $O(K_m \times N \times \log N)$ 。

4 仿真结果

4.1 系统仿真场景

为了简便，将多个子载波捆绑在一起组成 RB(Resource Block)进行分配，系统的详细参数请见表 3。

表 3 系统参数

参数	取值
载频	2 GHz
系统带宽	10MHz
RB	50

参数	取值
小区数	7
小区半径	500m
用户分布	均匀
基站总发射功率	20 W
噪声功率谱密度	-174dBm/Hz
阴影衰落方差	8dB
大尺度衰落	$L=128.1+37.6\log_{10}(R)$
多径模型	Typical Urban

4.2 性能评估

仿真将所提算法与文献[3]~[5]中的软频率复用方案进行比较,主要比较了系统频谱效率、不能满足最小速率需求的用户数以及小区边缘的平均速率。

从图 1 可以看出,在用户数较少的情况下,随着用户数的递增,由于用户数的增加所带来的多用户分集增益的增加,所提方案、软频率复用方案 1 以及软频率复用方案 3 的系统频谱效率也在提高,但是当用户数继续增多,这三种方案的系统频谱效率都会随着用户数的增多而下降,这是因为为了满足更多用户的最小速率需求,性能好的用户所能占用的子载波资源就越来越少,相应的系统的频谱效率也会降低。

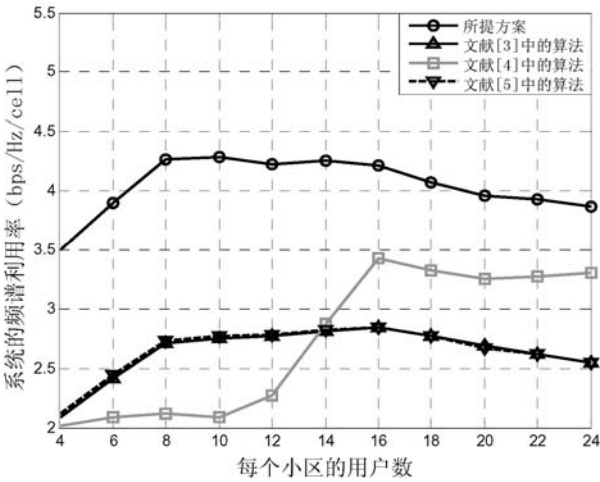


图 1 系统频谱效率

从图 2 中可以看出,随着用户数的增多,不能满足最小速率需求的用户数也随着增多,特别是软频率复用方案 1 和软频率复用方案 3,而所提方案和软频率复用方案 2 几乎能够满足所有用户的最小速率需求,这是因为在软频率复用方案 1 和软频率复用方案 3 中,每个小区可用的子载波资源受限,从而随着用户数增多,不能满足最小速率需求的用户数会急剧增多。

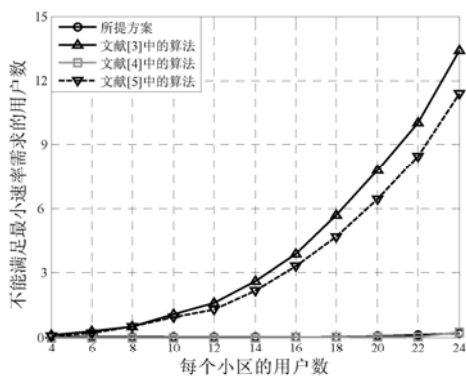


图2 不能满足最小速率需求的用户数

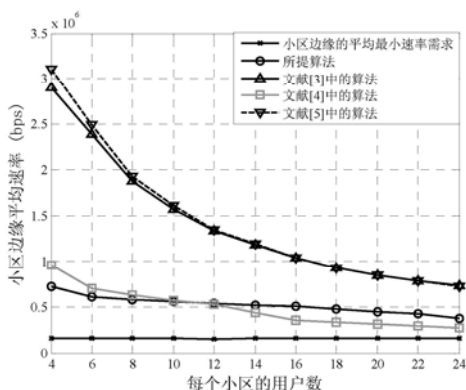


图3 小区边缘平均速率

从图3可以看出，每种方案都能满足边缘用户的平均最小速率需求。

5 结论

在这篇文章中，我们提出了一种多小区 OFDMA 系统的两级分布式子载波分配方案。首先将子载波分配问题分解为小区间子载波分配和小区内子载波分配两个子问题，然后从三个方面学习了所提的算法：系统频谱效率、不能满足最小速率的用户数以及小区边缘平均速率。仿真结果证明了所提算法可以有效的提高系统频谱效率，而且几乎所有用户的最小速率需求能够得到满足，尽管小区边缘平均速率较低，但是可以满足小区边缘的平均最小速率需求。进一步的工作将会集中研究所提算法在其它网络中的应用以及更多分布式资源分配算法。

参 考 文 献

- [1] J. Jang and K. B. Lee, "Transmit Power Adaptation for Multiuser OFDM Systems", IEEE JSAC, vol. 21, no.2, pp.171-178, Feb. 2003.
- [2] D. Tse and S. Hanly, "Multi-access fading channels: part I: polymatroidal structure, optimal resource allocation and throughput capacities," IEEE Trans. Inform. Theory, vol. 44, no. 2, pp. 2796-2815, Nov. 1998.

- [3] 3GPP TSG-RAN WG1 R1-050738, “Interference mitigation–Considerations and Results on Frequency Reuse”, Siemens, 29 August–2 September, 2005.
- [4] 3GPP TSG-RAN WG1 R1-050833, “Interference mitigation in evolved UTRA/UTRAN”, LG Electronics, August, 29 – September, 2, 2005.
- [5] 3GPP TSG-RAN WG1 R1-051059, “Inter-Cell Interference Mitigation for EUTRA”, Texas Instruments, 10 October - 14 October, 2005.
- [6] Deng Gang, Zou Ting, Xu Ning, Zhang Ping, “A Downlink Radio Resource Allocation Algorithm Based on Inter-Cell Interference Mitigation for Multi-Cell OFDMA System”, Communications and Networking in China, 2006.
- [7] Guoqing Li and Hui Liu, “Downlink Radio Resource Allocation for Multi-Cell OFDMA System”, IEEE Transactions on Wireless Communications, Vol. 5, No. 12, DEC 2006.
- [8] M. Pischella and J. C. Belfiore, “Distributed Resource Allocation for Rate-Constrained users in Multi-Cell OFDMA Networks”, IEEE Communications Letters, VOL. 12, NO. 4, APRIL 2008.
- [9] A. J. Goldsmith and S.-G. Chua, “Variable-rate variable-power MQAM for fading channels,” IEEE Transaction on Communications vol. 45, pp. 1218–1230, Oct. 1997.
- [10] B. Korte, and J. Vygen, *Combinatorial Optimization: Theory and Algorithms, Second Edition*. Springer Verlag, 2002.
- [11] C. H. Papadimitriou and K. Steiglitz, *Combinatorial Optimization: Algorithms and Complexity*. Prentice-Hall, 1982.

作者简介:

邱涛, 男(1986-) 硕士研究生 无线通信系统中的资源分配

贺志强, 男(1975-) 副教授 信息理论与移动通信

牛凯, 男(1976-) 副教授 信息理论与移动通信

田宝玉, 男 1946 教授 数字移动通信理论与技术

一种下行单用户协作多点传输的调度算法

张成强 贺志强 牛 凯 田宝玉

(北京邮电大学, 北京 100876)

摘 要: 本文提出了一种用于小区间干扰协调的协作多点 (Coordinated Multi-Point, CoMP) 调度算法, 包括特定 UE 选择协作小区选择和协作小区间的联合调度。该算法利用多根发射天线为特定用户联合传输, 小区间干扰变成有用信号。仿真结果显示所提方案在可接受的调度复杂度下大幅提高小区吞吐量; 与软频率复用方案相比, 协作多点传输方案更有效、更公平。
关键词: 通信技术; 协作多点; 频域划分; 软频率复用

An Scheduling Algorithm of Downlink Coordinated Multi-Point Transmission/Reception for Single-User

Zhang Cheng qiang He Zhi qiang Niu Kai Tian Bao Yu

(Beijing University of Posts and Telecommunications, Beijing 100876)

Abstract: In order to coordinate the inter-cell interference, a Coordinated Multi-Point (CoMP) scheduling algorithm was proposed, which contained selection of coordinated cells and joint scheduling between coordinated cells. A specific user received signals from multiple antennas in this algorithm, so the main inter-cell interference became signals. Simulation results display that this scheme had a large cell throughput with an accepted level of scheduling complexity; this algorithm got higher efficiency and more fairness than soft-frequency reuse scheme.

Keywords: Communication technology; Coordinated Multi-Point; frequency partition, soft-frequency reuse

1 介绍

协作多点 (Coordinated Multi-Point, CoMP) 传输技术[1]能够有效地利用小区间干扰, 提高小区边缘以及系统平均吞吐量, 并扩大覆盖率。而在协作多点传输中, 单个用户可能接收多个 eNB 同时发送的数据, 来提高接收信号质量或消除干扰。

如果系统所有 eNB 共同协作进行传输, 将提供最可观的小区边缘和平均吞吐量增益。然而, 所有 eNB 的数据和 CSI 共享对系统回程链路提出了极高的要求, 复杂度高至难以实现。如何在可接受复杂度及回程容量要求下, 实现 CoMP 传输成了关键问题。

为了简化叙述, 本文约定^[2]:

CoMP 传输集 (CoMP transmission point/points): 参与 PDSCH 传输的 eNB 集合; CoMP 报告集 (CoMP reporting set): 用户进行信道质量测量的小区集合。在本文中, CoMP 报告集

等同于 CoMP 协作集，即不同 CoMP 报告集的 eNB 不会相互协调进行通信。在一些提案中，为特定用户选择 CoMP 传输集的过程被称为小区成簇（cell clustering）^{[3][4][5]}。

本文第二部分描述所提的 CoMP 传输集选择方案，第三部分给出仿真结果及分析，第四部分总结全文。

2 CoMP传输集选择

目前存在三种 CoMP 传输集选择方案：纯 UE-Specific 选择方案，固定划分方案，及折中的小区成簇方案^{[3][4]}。在纯 UE-Specific 选择方案中每个用户都有独立的 CoMP 传输集，不同用户的 CoMP 传输集可能相互重叠，需要联合调度，回程信息量大，实现复杂度高；在固定划分方案中网络预先划分 CoMP 传输集，CoMP 传输集为位于此地理空间上的用户提供服务，实现简单，增益有限；而折中的小区成簇方案是预先划分 CoMP 报告集，根据用户反馈信息从 CoMP 报告集中选择 CoMP 传输集，位于 CoMP 报告集中心的用户提供较高的 CoMP 传输增益，位于 CoMP 报告集边缘的用户性能增益不明显。

在上述方案中，纯 UE-Specific 选择方案提供的吞吐增益最大，但需要全网统一协调，原因是不同 CoMP 传输集有重叠。如果能够分离出各用户的 CoMP 传输集，那么实现复杂度将会大幅降低。固定划分方式虽然将 CoMP 传输集在空间上分开，但是损失了空间分集增益。本文提出一种基于频域划分的 CoMP 传输集选择方案，允许在空间上相互重叠。避免了全网协调，几乎不损失空间分集增益，并且大幅降低复杂度。仿真结果显示，此方案获得的吞吐量接近于纯 UE-Specific 选择方案。

在 UE-Specific 选择方案中，CoMP 传输集选择由用户信道状态决定，是动态变化的，难以进行频域分离。不过，CoMP 传输集是 CoMP 报告集的子集；CoMP 报告集由用户的服务小区静态/半静态确定。从频域上分离 CoMP 传输集的问题可以简化为从频域上分离 CoMP 报告集的问题。

在 LTE 系统中，用户的主要干扰源来自与服务小区最近的小区集合。假定服务小区内的用户 CoMP 报告集是服务小区与周围六个小区并集。以图 1 为例，小区 0 的用户 CoMP 报告集是小区 0、1、2、3、4、5、6 在灰白色频段上的集合；小区 1 的用户 CoMP 报告集是小区 0、1、2、6、7、8、9 在灰色频段上的集合，等等。

具体而言，将各小区中可用 CoMP 频段等分成三份：一份为发起 CoMP 频段，两份为协作 CoMP 频段。本小区发起 CoMP 传输时，使用发起 CoMP 频段；本小区被请求进行 CoMP 传输时，使用协作 CoMP 频段。以 19 个小区（见图 1）为例，小区 0、7、9、11、13、15、17 的可用 CoMP 频段可划分如下：

其他小区可 CoMP 频段划分据此类推。当小区 0 发起 CoMP 传输时，将使用灰白色频段，而 CoMP 传输集中的其他 eNB 也使用灰白色频段辅助服务小区 0 进行 CoMP 传输；当小区 1 发起 CoMP 传输时，将使用灰色频段，而 CoMP 传输集中的其他 eNB 也使用灰色频段辅助服务小区 1 进行 CoMP 传输；等等。

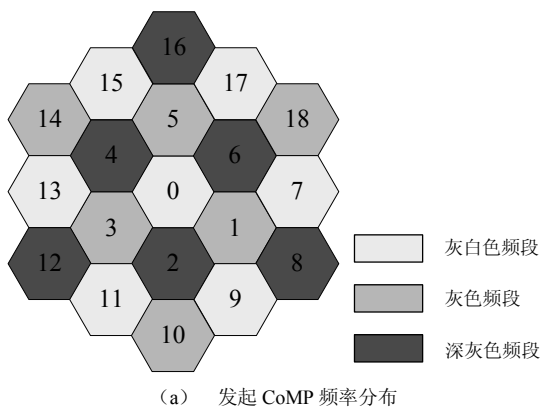


图 1

3 系统级仿真及分析

本文所提的协作多点通信的调度算法，在 CoMP 传输集的选择上有别于现有方案；同时，这种基于频域划分的 CoMP 传输集选择方案与软频率复用方式有相通之处。为了进一步验证所提的协作多点通信的实现机制的性能，本文引入仿真进行分析。

图 2 是几种 CoMP 传输集选择方案用户 SINR 的累积分布图。仿真场景具体参数见附录。如果用户使用 CoMP 传输，“频谱效率”有所提高，则进行 CoMP 传输，称其为“CoMP 用户”，否则称为“非 CoMP 用户”。“频谱效率”定义为可获速率与 CoMP 传输集中 eNB 数目的比值。

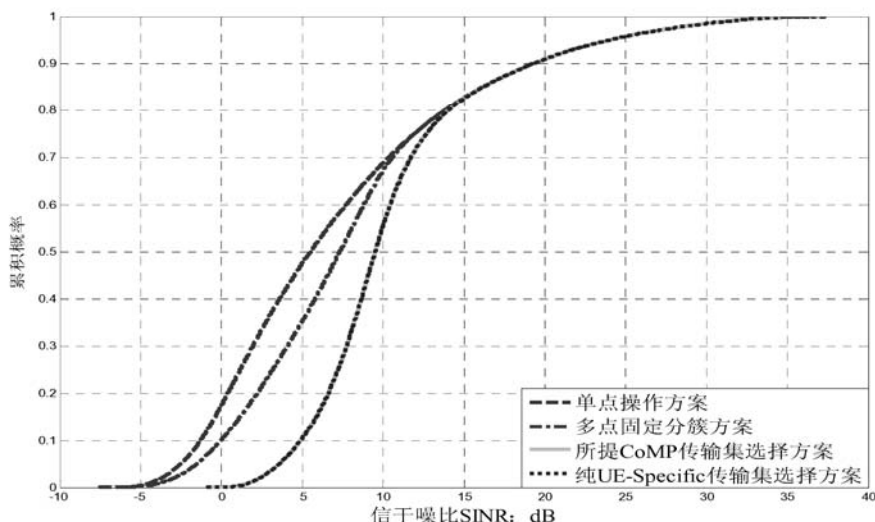


图 2 不同 CoMP 传输集选择方案下的 SINR 累积分布图

在 CoMP 用户的 SINR 中，CoMP 传输集中的原干扰成了有用信号^[6]。从图 2 中可以看出，

所提方案能够获得几乎与纯 UE-Specific 方式相同的性能，而远优于固定方式。在所提的基于频域划分的选择方案下，服务小区可以自主地为本小区 CoMP 用户选择传输集、调度资源，也无需与其他节点协调，其实现复杂度与固定方式的复杂度相仿。

LTE 系统采用“软频率复用”技术进行小区间干扰协调^[7]，减轻小区间干扰。本文以提案[7]的软频率复用方案为参考，即边缘用户频率复用因子为 3，中心用户频率复用因子为 1，与所提协作多点传输方式进行性能比较。

仿真场景：10MHz 带宽分成 50 个 RB（Resource Block），具体参数见附录。在仿真过程中，采用等功率分配的比例公平算法调度 RB。两个参考方案分别是传统蜂窝网络传输方案；软频率复用方案。在仿真过程中，假设软频率复用边缘用户与多点协作传输 CoMP 用户的判定准则是一致的；公平系数被定义为中心用户平均速率与 CoMP/边缘用户平均速率比；以仙依容量计算链路速率，并计算系统平均频谱效率。在不同的 CoMP 频段资源（或复用因子为 3 的频段资源）分配下，系统的公平系数和平均频谱效率变化如图 3 所示。

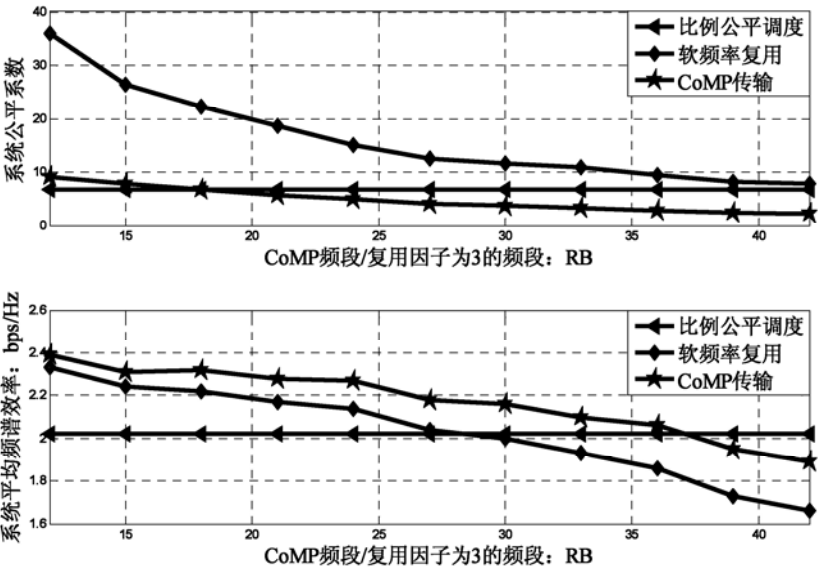


图 3 CoMP 传输方案与软频率复用方案比较

- (1) 总体趋势显示，CoMP 频段的资源越多，系统公平系数越低，即公平性越好，但是系统的平均频谱效率下降。原因在于，CoMP 频段越多，CoMP 用户获得的速率越高；由于所有用户共用有限频谱资源，地理位置比较好的非 CoMP 用户资源减少，它们的平均速率逐渐下降。
- (2) 与软频率复用方案相比，CoMP 传输方案公平系数一直低于软频率复用方案；而有效性一直高于软频率服用方案。这种增益是由 CoMP 传输模式下，多个 eNB 为某个用户服务提供的。CoMP 传输模式的代价在于需要多点精确同步。

4 总结

本文提出了一种基于频域划分的 UE-Specific CoMP 传输集的选择方案，和相应的协作调

度算法。在所提方案中，相邻小区 CoMP 传输集在频域上隔离，服务小区 eNB 可以通过用户反馈信道信息动态地为用户自主选择 CoMP 传输集，无需全网统一协调，实现复杂度低。仿真结果显示：（一）所提方案以趋近固定划分方式的调度复杂度，获得了接近纯 UE-Specific 选择方案的小区吞吐量；（二）所提 CoMP 传输方案在公平性和有效性方面都优于软频率复用方案。

参 考 文 献

[1] 3GPP, R1-084294, “Efficient HARQ Protocol for SIC based DL CoMP”, Fujitsu
[2] 3GPP, R1-092199, “Draft Report of 3GPP TSG RAN WG1 #56bis” v0.2.0
[3] 3GPP, R1-091377, “Updates on Cell Clustering for CoMP Transmission Reception”, Nortel
[4] 3GPP, R1-090725, “Setup of CoMP cooperation areas”, Nokia Siemens Networks, Nokia
[5] 3GPP, R1-084464, “Cell Clustering for CoMP Transmission/Reception”, Nortel
[6] 3GPP, R1-090613, "Discussions on CoMP SU-MIMO", Samsung
[7] 3GPP, R1-050738, “Interference mitigation – Considerations and Results on Frequency Reuse”, Siemens

附录：仿真参数

参数	参数选择
站间距（ISD），网络拓扑	500m，正六边形，7 小区+ wrap
用户分布、数量	随机均匀分布，15 用户/小区
带宽/资源块数量	10MHz/50
天线总发射功率/接收端噪声系数	40dBm/5dB
对数正态阴影效应分布的标准方差	8 dB
用户移动速度	3Km/h
业务模型	Full buffer

作者简介：

张成强（1985-），男，在读硕士研究生，北京邮电大学，主要研究方向移动通信；
贺志强（1975-），男，北京邮电大学，副教授，主要研究方向信息理论与移动通信
牛凯（1976-），男，北京邮电大学，副教授，主要研究方向信息理论与移动通信
田宝玉（1946-），男，北京邮电大学，教授，主要研究方向数字移动通信理论与技术

一种基于图码的量子稳定子码构造方法的研究

朱修利 赵生妹

(南京邮电大学 通信与信息工程学院, 南京 210003)

摘要: 本文给出一种从图码构造量子校验矩阵的方法, 这种构造方法比基于计算机穷举法和群论方法构造的量子码更加简单。通过对该方法获得的(24,6)量子码和(32,8)码性能分析和比较, 可以发现这种量子码和 Thomas 的性能相似, 这是一种有效的量子码构造方法。

关键词: 量子码; 图码; 稳定子码; 稳定子生成元

A Construction of quantum code based on Graph code

Zhu Xiu li Zhao Sheng mei

(Institute of Signal Processing and Transmission, Nanjing University of Posts and Telecommunications .Nanjing 210003)

Abstract: In this paper we gave a construction of quantum codes derived from Graph codes. This construction method is simpler than those based on Classical Linear Codes and theory of groups. By the simulation, the decoding performance of (24,6) quantum code and (32,8) quantum code constructed by this method is similar to the results discussed by Thomas in his paper.

Keywords: Quantum code; graph code; stabilizer code; stabilizer generators

1 引言

迄今为止, 量子纠错理论日趋完善, 几乎所有经典纠错编码方案都已经被移植到量子领域中^[1]。Calderbank、Shor 以及 Steane 提出了一个从现有线性分组码构造量子纠错码的方法^{[2][3]}, 简称 CSS 码。Chau 利用经典卷积码的方法构造出量子卷积码^[4]; David J.C. MacKay 等人于 2004 年在 CSS 码基础上, 利用循环稀疏序列, 即满足自对偶条件, 提出了四种构造校验矩阵的方法, 并在此基础上构造量子码^[5]; 张权根据经典 Turbo 原理构造出量子 Turbo 码^[6]等。由于大部分经典纠错码都是图码, 因而可以从图论的角度进行入手, 直观地进行量子纠错码的构造。本文从图码的基本理论出发, 给出从图码到稳定子码的构造过程, 并以实例进行分析和比较。

2 图码和稳定子码

一个图可以表示为一个集合 $G = (V, E)$, 其中 $V = \{v_0, v_1, \dots, v_{n-1}\}$ 是 n 个顶点(节点)的集合, 而 E 是集合 V 中一对元素的集, $E \subseteq V \times V$, 顶点对 $\{v_i, v_j\} \in E$ 被称为边。若仅考虑

无向图，即连接各顶点的边是没有方向的。而且所研究的图是简单的不包含自环的图，即 $\{v_i, v_i\} \notin E$ 。满足 $V' \subseteq V$ 和 $E' \subseteq E$ 条件的图 $G' = (V', E')$ 是 G 的子图 ($G' \subseteq G$)。给出顶点的子集为 $A \subseteq V$ ，则子图 $G(A) \subseteq G$ 包含了顶点集 A 以及边集 $\{\{v_i, v_j\} \in E \mid v_i, v_j \in A\}$ ， E 中边所连接的顶点都在集合 A 中。图 G 中除去子图余下的图 \bar{G} 包含了顶点集为 $\bar{V} = V$ 和边集为 $\bar{E} = V \times V - E$ 。两个同构图在结构上是等价的。如果两个图 $G = (V, E)$ 和 $G' = (V', E')$ 是同构的，则肯定会在 V 中存在一个映射 π 使得 $\{v_i, v_j\} \in E \Leftrightarrow \{\pi(v_i), \pi(v_j)\} \in E'$ 。

单量子位和环境相互作用可用 $\{I, X, Y, Z\}$ 四个算子表示，称为 Pauli 算子。n 量子位的 Pauli 算子是单量子位 Pauli 算子的 n 重直积：

$$M = X_1 \otimes Z_2 \otimes I_3 \otimes Y_4 \otimes \cdots \otimes X_n \quad (1)$$

令 S 为 n 量子位 pauli 算子群 G_n 的一个 Abel 子群。由于 S 中的元素都相互对易，他们在 n 量子位的 Hilbert 空间中可以同时对角化。记 S 中所有元素本征值为 +1 的共同本征空间为 H_s ，即当且仅当对所有 $M \in S$ ，有：

$$M|\varphi\rangle = |\varphi\rangle \quad (2)$$

成立，才有 $|\varphi\rangle \in H_s$ 。如果一个量子码的码空间就是 H_s ，则称这个量子码是稳定子码[8]，并称子群 S 是这个码的稳定子。

稳定子 S 的生成元可以校验稳定子码。由于 S 的生成元 M_1, M_2, \cdots, M_k 是一组相互对易的厄米算子，它们可以同时取确定值，因此可以同时对这组算子的集体测量诊断出错误。如果编码态没有错误出现，那么对每个生成元 M_i ，测量结果均为其本征值 +1。如果对某个 M_i 测量结果为 -1，表明错误出现。由于任意出错都可用 G_n 群元算子的和表示，一个特定的错误 E_a 和稳定子的一个生成元 M_i 或对易或反对易。如果 E_a 和 M_i 对易，那么

$$M_i E_a |\varphi\rangle = E_a M_i |\varphi\rangle = E_a |\varphi\rangle \quad (3)$$

其中 $|\varphi\rangle \in H_s$ ，这时出错态仍保留在码空间中。如果 E_a 和 M 反对易，则有

$$M_i E_a |\varphi\rangle = -E_a M_i |\varphi\rangle = -E_a |\varphi\rangle \quad (4)$$

出错将把编码态变换到码空间以外。总结上面两种情况，对于稳定子生成元 M_i 和出错 E_a ，可以写出

$$M_i E_a |\varphi\rangle = (-1)^{S_{ia}} E_a |\varphi\rangle \quad (5)$$

其中当 E_a 和 M_i 对易时， $S_{ia} = 0$ ；当 E_a 和 M_i 反对易时， $S_{ia} = 1$ 。所以对出错态 $E_a |\varphi\rangle$ ，测量稳定子 S 的一组生成元 $\{M_i\}$ ，将得到一组本征值： $(-1)^{S_{ia}} (i = 1, 2, \cdots, n - k)$ 。式中的 S_{ia} 就是错误 E_a 的指错子。对于非简并码，不同的 E_a ，指错子 S_{ia} 不同，从而测量稳定子的 n-k 个生成元，可以纠正所出错的码^[9]。

3 基于图码的量子稳定子码构造方法

图码是一种 GF(4)域上的自对偶加性码^[7]，它的邻接矩阵可写为如下形式：

$$\Gamma = \begin{pmatrix} M_x & B \\ B^T & M_y \end{pmatrix} \quad (6)$$

通过理论推导，能够得到一个稳定子码，其校验矩阵为：

$$H = D \cdot (I | M_y) \quad (7)$$

式中 M_x ， M_y 为对称矩阵， B^T 是 B 的转置， D 为 B 的正交矩阵。下面我们以图 1 为例说明构造量子稳定子元的过程：

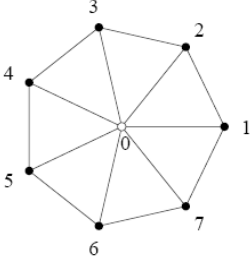


图 1 clique 图

$$\Gamma_{W_7} = \left(\begin{array}{c|cccccccc} 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ \hline 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \end{array} \right)$$

图 2 邻接矩阵

因为图可以表示为邻接矩阵 Γ ，这是一个 $|V| \times |V|$ 的矩阵，其中如果 $\{v_i, v_j\} \in E$ ，既 v_i ， v_j 可被一条边连接。则 $\Gamma_{i,j} = 1$ ，反之 $\Gamma_{i,j} = 0$ 。对于简单图来讲，矩阵对角线全部为 0，即 $\Gamma_{i,i} = 0$ 。而且一个无向图的邻接矩阵是对称的，即 $\Gamma_{i,j} = \Gamma_{j,i}$ 。

因此可以由图 1 获得这个图的的邻接矩阵 Γ_{W_7} 如图 2。

从图 2 中可以看出 B 为 1×7 的矩阵，所以它的一个正交矩阵 D 为：

$$D = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix}, \text{ 由 } H = D \cdot (I | M_y) \text{ 可得量子校验矩阵:}$$

$$H = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 \end{pmatrix} \quad (8)$$

可以看出从校验矩阵 H 中任取两行表示为 $(\alpha | \beta)$ ， $(\alpha' | \beta')$ ，如第一、二两行则 $\alpha = (1000001)$ ， $\beta = (1100011)$ ， $\alpha' = (0100001)$ ， $\beta' = (0010010)$ ，其辛内积为 $\alpha \cdot \beta' \oplus \alpha' \cdot \beta = 0$ 。既任意两矢量正交，满足对易条件。校验矩阵 H 对应着稳定子元的矢量偶表示[8]，相应 GF(4)域的稳定子生成元为：

$$C = \begin{pmatrix} Y & X & I & I & I & X & Y \\ I & Z & X & I & I & X & X \\ X & X & Z & X & I & X & X \\ X & I & X & Z & X & X & X \\ X & I & I & X & Z & I & X \\ X & I & I & I & X & Y & Y \end{pmatrix} \quad (9)$$

这样我们就可以得到 6 组生成元: $M_1=YXIII XY$, $M_2=IZXII XX$, $M_3=XXZXIXX$, $M_4=XIXZXXX$, $M_5=XIIXZIX$, $M_6=XIIIXYY$

根据稳定子码的编码公式:

$$|c_1 \dots c_k \rangle_L = \left(\prod_{i=1}^{n-k} (I + M_i) \right) \overline{X}_1^{c_1} \dots \overline{X}_k^{c_k} |0 \dots 0 \rangle = \overline{X}_1^{c_1} \dots \overline{X}_k^{c_k} \left(\sum_{M \in S} M |0 \dots 0 \rangle \right) \quad (10)$$

可选择 $\overline{X} = XXXXXXXX$, 便可得到一位(7,1)量子稳定码。

4 仿真及性能分析

根据上面所述, 现通过这种方法构造不同码率、不同码长的量子码, 并分析其性能。由于复杂图的仿真较为复杂, 本文仅考虑简单图的量子码。传输信道的错误模型为退极化信道, 其中信息在信道传输中比特错误概率为 P , 即发生比特翻转错误、相位翻转错误以及比特和相位均发生错误的概率各为 $P/3$, 未出错的概率为 $1-P$, 采用稳定子码的译码方法进行译码, 结果如图 3 所示。图中, 纵坐标为误帧率, 横坐标为比特发生错误的概率。

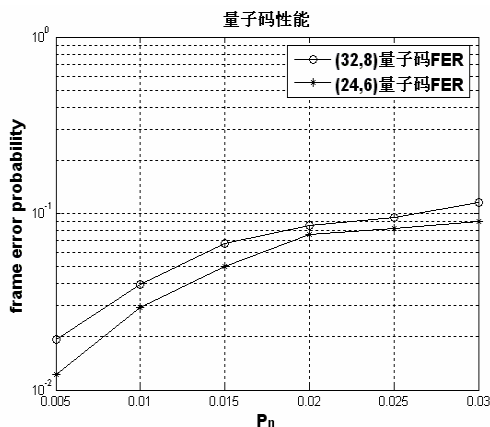


图 3 (24,6)和(32,8)量子码误帧率性能

这两种量子 LDPC 码的码率都为 $1/4$, 不过码长不同, 其中(24,6)码性能比(32,8)码稍好。从图中可以看出, 码长越短, 码字的性能越好, 在比特翻转概率为 0.005 处, 前者比后者要优 0.5dB 左右。当然, 这也不是绝对的, 因为码字的性能还跟码距, 短环的存在等其他因素有关。仿真结果与 Thomas^[10]等人所构造的量子码在退极化信道中的仿真结果相类似。

5 结束语

本文在图码的理论基础上,研究了用图码构造稳定子生成元,并以(24,6)和(32,8)量子码为例进行了仿真,对仿真结果进行了分析和比较,证明从图码构造量子稳定子码的可行性。此外,图码有很多,几乎所有的线性码都可以转化成图码,如何从众多图码中选出性能好的码字作为量子纠错码,可以是后续的研究方向。

参 考 文 献

[1] A.M.Steane,Error Correcting Codes in Quantum Theory[J].Phys.Rev.Letters,1996, 77(5): pp:793- 797.

[2] A.R.Calderbank and P.W.Shor, Good Quantum Error-Correction Codes Exist[J], Phys.Rev.A, 1996, 52, pp:2493

[3] Gottesman, Class of Quantum Error-Correcting Codes Saturating the Quantum Ham-ming Bound[J], Phys.Rev.A, 1996, 54(3), pp:1862-1869

[4] H.F.Chau, Quantum Convolutional Error Correction Codes[J], Quantum Computing and Quantum Communication(QCQC)'98, 1998, pp:314-324

[5] D.Mackay, G.Mitchison and P.Mcfadden, Sparse Graph Codes for Quantum Error C-orrection[J], IEEE Trans.Inform.Theory, Oct.2004, 50: pp:2315–2330. Also arXiv/qua-nt-ph/0304161

[6] 唐朝京,高峰, 量子 Turbo 码[J], 物理学报, 2002, 51(1), pp:15-19

[7] Markus Grassl, Andreas Klappenecker, and Martin Rotteler, Graphs, Quadratic Forms, and Quantum Codes [EB/OL], <http://lanl.arxiv.org/abs/quant-ph/0703112v1>, Mar 2007

[8] Gottesman, Stabilizer Codes and Quantum Error Correction[EB/OL], [http:// lanl.arxiv.org/abs/quant-ph/9705052v1](http://lanl.arxiv.org/abs/quant-ph/9705052v1), May 28,1997

[9] M.A.Nielsen and I.L.Chuang, Quantum Computation and Quantum Information[M].C- amberidge university press,Cambridge, 2000

[10] Thomas.Camara, H.Ollivier, J.P.Tillich, Constructions and Performance of Classes of Quantum LDPC Codes[EB/OL], <http://lanl.arxiv.org/abs/quant-ph/0502086>, Feb.2005

作者简介:

朱修利 (1984-), 男, 江苏宿迁, 南京邮电大学通信与信息工程学院研究生, 主要研究方向: 量子编码

赵生妹(1968-), 女, 江苏丹徒人, 南京邮电大学教授, 主要研究方向: 量子通信与信息处理, 无线通信与信号处理

附联系方式:

姓名: 朱修利

通信地址: 江苏省南京市新模范马路 66 号南京邮电大学 103 信箱

邮编: 210003

电子邮件: Y070628@njupt.edu.cn

基于SFBC-OFDM系统的时变信道估计和信号检测联合算法

任术波 郭俊奇 项海格

(北京大学 信息科技学院 卫星与无线通信实验室, 北京 100871)

摘要: 在时变多径信道环境下, 针对导频辅助的无循环前缀 SFBC-OFDM 系统提出了一种信道估计和信号检测联合处理的单块迭代和多块递归的干扰抑制算法, 仿真验证表明: 该算法不仅可以实现时变信道环境下系统良好的信道估计与跟踪, 还可以获得优良的数据符号检测性能, 为这一系统在高速移动环境下的宽带无线通信数据可靠传输提供了理论基础。

关键词: 信道估计; 信号检测; 单块迭代; 多块递归

A Joint Algorithm of Time-varying Channel Estimation and Signal Detection for SFBC-OFDM Systems

Ren Shu bo Guo Jun qi Xiang Hai ge

(Satellite and Wireless Communication Lab, EECS, Peking University, Beijing, 100871)

Abstract: An interference-cancellation algorithm based on the joint processing of time-varying channel estimation and signal detection in the form of single-block iteration and multiple-block recursion is proposed for non-cyclic prefix space-frequency block coded-orthogonal frequency division multiplexing (non-CP SFBC-OFDM) systems. Simulation results show that the proposed algorithm achieves not only good time-varying channel estimation/tracing but also good data detection performance. Therefore, it provides a solid theoretical foundation for broadband wireless communications of this system in high-speed mobile environments.

Keywords: Channel estimation; Signal detection; Single-block iteration; Multiple-block recursion

1 引言

随着蜂窝移动通信、无线互联网和无线多媒体等宽带数据业务传输需求的快速增长, 人们对未来新一代宽带无线通信的传输速率和质量赋予了更高的期望, 而多输入多输出(MIMO)的空间分集被认为是实现无线通信数据稳定可靠传输的关键技术, 它与正交频分复用(OFDM)技术相结合而构成的各种 MIMO-OFDM 系统^[1-6]不仅能确保无线通信传输质量, 还可以克服宽带信道的频率选择性衰落, 成为新一代无线通信系统最具竞争力的备选方案, 因此相关各系统接收机的信道估计和信号检测联合处理等关键技术成为人们研究的热点。

目前时变信道环境下实现无线宽带数据的高速移动接收已经成为未来新一代移动通信发

展的主要方向，而最大似然信道估计和信号检测联合处理技术则是解决这一问题的经典接收机信号处理算法，但计算复杂度高不易于实现，因此寻找次最优的各种联合处理算法成为人们研究的重点。文献[7][8]采用期望最大化（EM）算法来降低最大似然处理技术的计算复杂度实现快速时变信道估计和信号检测的联合处理。而文献[9]则是在缓慢块衰落信道环境下把最大似然的联合处理技术转换为球形解码来实现，在降低实现复杂度的同时获得比较好的系统误码率性能。

与上面提到的信道估计和信号检测联合处理算法思路不同，本文以文献[10]给出的导频和数据交替发送且均无保护间隔的无循环前缀 SFBC-OFDM 系统为研究背景，提出了一种相对简单的时变信道环境下信道估计和信号检测联合处理的单块迭代和多块递归的干扰抑制算法：单块迭代干扰抑制算法可以实现每一个接收信号块信道估计和信号检测的快速收敛，使系统性能快速趋于稳定状态，而多块递归干扰抑制算法则可以进一步获得当前每个接收信号块更精确的信道估值，在实现系统良好信道估计与跟踪的同时大大提高了系统的信号检测性能。该算法为这一无循环前缀 SFBC-OFDM 系统实现快速时变信道环境下的高速无线宽带数据接收提供了强大的保障。

接下来分 4 部分展开阐述：第 2 节简单介绍了基于导频辅助的无循环前缀 SFBC-OFDM 系统模型；第 3 节重点介绍了时变信道环境下信道估计和信号检测的联合处理算法；第 4 节给出仿真结果和性能分析；最后给出本文结论。

2 系统模型

文献[10]对导频序列和信息数据块交替发送且均无保护间隔的无循环前缀 SFBC-OFDM 系统进行了深入的研究，提出了一套切实可行的发射机和接收机模型，下面分别予以简单介绍。

2.1 发射机模型

在基于导频辅助的无循环前缀 SFBC-OFDM 系统中（以 2×1 发射接收天线系统为例），发射端两根发射天线每相邻两个频域子载波位置上的发射 OFDM 信息数据符号采用 Alamouti 结构的空频分组编码，则第 k 个信号块传输周期内，发射天线 1 和发射天线 2 上 OFDM 信息数据块和导频序列所构成的无保护间隔发射信号块结构如图 1 所示，分别记为 \mathbf{x}_1^k 和 \mathbf{x}_2^k 。

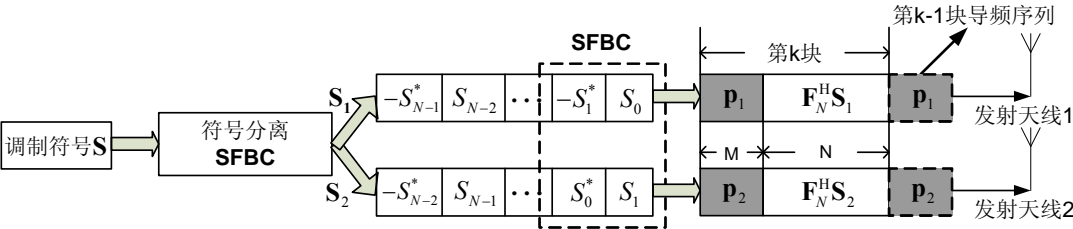


图 1 导频辅助的无循环前缀 SFBC-OFDM 发射机模型

2.2 接收机模型

针对图 1 的发射机模型, 图 2 给出了基于导频辅助的无循环前缀 SFBC-OFDM 系统的接收联合处理流程。

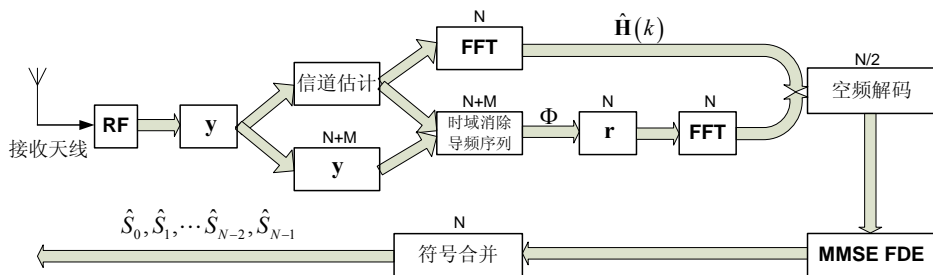


图 2 导频辅助的无循环前缀 SFBC-OFDM 接收机模型

在静态或时变多径信道环境下，两根发射天线上第 k 时隙内的发射信号块 \mathbf{x}_1^k 和 \mathbf{x}_2^k 分别经过多径衰落信道 $\mathbf{h}_{1,k}$ 和 $\mathbf{h}_{2,k}$ （由于采用一根接收天线，因此省略了接收天线脚标）到达接收端后，利用信道估计和信号检测的联合处理算法（下一节详细描述）即导频和数据相互影响的干扰抑制技术，一方面在时域上消除接收 OFDM 信息数据符号对接收导频序列的干扰来估计各发射天线到接收天线的多径信道冲激响应；另一方面在时域上消除接收导频序列对接收 OFDM 信息数据符号的干扰，然后进行相应信号处理操作（主要包括“剪切相加”操作、FFT 变换、Alamouti 空频解码、“准”MMSE 频域均衡等信号处理）检测接收数据符号，直到无线数据业务传输完毕为止。

3 时变信道环境下信道估计和信号检测的联合处理算法

由于多径信道的时延扩展, 导频序列和 OFDM 信息数据符号交替发送且均无保护间隔的这一无间隙发射信号块结构, 会使接收导频序列和接收 OFDM 符号在相邻处交叠在一起产生相互干扰, 图 3 给出了形象的描述: 灰色格型三角形表示接收导频序列因信道多径时延扩展而产生的“拖尾”, 它对接收 OFDM 信息数据符号产生干扰, 大大降低了信息数据符号的检测性能; 而白色格型三角形表示接收 OFDM 信息数据符号因信道多径时延扩展而产生的“拖尾”, 它对接收导频序列产生干扰, 严重影响了多径信道的估计性能。此外在高速移动的时变多径信道环境下, 多普勒频率扩散的存在会进一步增加多径信道冲激响应估计的难度, 系统跟踪信道的时变性能相应会越来越差, 从而导频序列和 OFDM 信息数据符号之间的相互干扰也会越来越难以消除, 因此大大影响了信息数据符号的检测性能, 随之会导致整个无线通信系统数据传输质量的严重下降, 甚至有可能出现无线通信传输中断等现象。

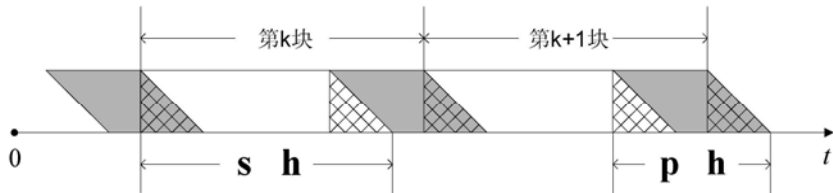


图 3 接收导频序列和接收 OFDM 符号的相互干扰示意图

为了解决上述问题，这里提出了一种时变信道环境下单块迭代和多块递归相结合的信道估计和信号检测联合处理算法即导频和数据相互干扰抑制算法。为了不失一般性，下面给出时变多径信道下无循环前缀 SFBC-OFDM 系统中任意第 k 个接收信号块的信道估计和信号检测联合处理流程（第一块和其它接收数据块处理基本一样）：

1) 首先利用第 $k-1$ 接收信号块接收处理所获得的信道估值 $\hat{\mathbf{h}}_{1,k-1}$ 和 $\hat{\mathbf{h}}_{2,k-1}$ 和两根发射天线上的发送导频序列 \mathbf{p}_1 和 \mathbf{p}_2 重构第 $k-1$ 块接收导频，然后消除其对当前第 k 接收信号块的干扰（假定干扰消除“干净”），并把消除干扰的当前第 k 接收信号块与第 $k+1$ 接收信号块联合看作一个“接收处理单元”，如图 4 所示；

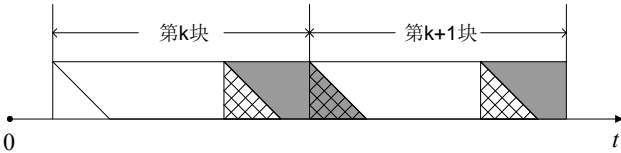


图 4 消除第 $k-1$ 块导频序列干扰的接收处理单元

2) 接着利用当前第 k 接收信号块接收处理获得的信道估值 $\hat{\mathbf{h}}_{1,k}$ 和 $\hat{\mathbf{h}}_{2,k}$ （当前块内第一次迭代采用第 $k-1$ 接收信号块获得的信道估值 $\hat{\mathbf{h}}_{1,k-1}$ 和 $\hat{\mathbf{h}}_{2,k-1}$ 作为当前第 k 接收信号块的信道估值）以及当前第 k 、第 $k+1$ 块发射导频序列 \mathbf{p}_1 和 \mathbf{p}_2 分别重构两块的接收导频，然后分别消除其对第 k 和第 $k+1$ 接收 OFDM 信息数据符号的干扰，如图 5 所示；

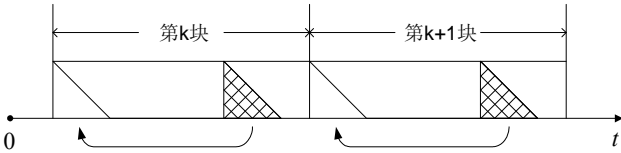


图 5 消除第 k 和第 $k+1$ 时隙导频序列干扰的接收处理单元

3) 其次对消除导频干扰的当前第 k 和第 $k+1$ 接收 OFDM 信息数据符号分别进行“剪切相加”操作、FFT 变换、Alamouti 空频解码、“准”MMSE 频域均衡以及符号判决，并重新符号调制得到第 k 和第 $k+1$ 两个时隙各自原始发送的 OFDM 信息数据符号的估值；

4) 接下来再利用当前第 k 接收信号块接收处理获得的信道估值 $\hat{\mathbf{h}}_{1,k}$ 和 $\hat{\mathbf{h}}_{2,k}$ 以及第 3) 步获得的两个时隙 OFDM 信息数据符号估值重构当前第 k 和第 $k+1$ 接收 OFDM 信息数据符号，并从图 4 所示的原始接收数据块中消除其对当前第 k 时隙内接收导频序列的干扰，如图 6 所示；

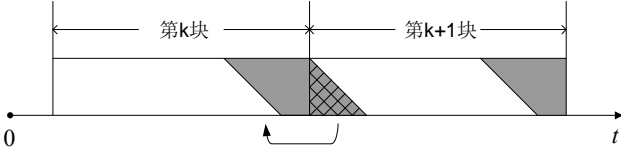


图 6 消除第 k 和第 $k+1$ 时隙 OFDM 符号干扰的接收处理单元

5) 再次利用消除第 k 和第 $k+1$ 接收 OFDM 信息数据符号干扰、比较“干净”的当前第 k 块

接收导频序列估计当前接收信号块内各个发射接收天线对之间的信道冲激响应估值（时变信道下为了防止出现性能平台不进行估计后处理操作），随后各天线对之间的信道估值再分别与各自递归一定数目的历史信道估值信息求平均作为当前第 k 块的信道估值 $\hat{\mathbf{h}}_{1,k}$ 和 $\hat{\mathbf{h}}_{2,k}$ ，然后跳转到第2)步循环执行，依次实现若干次迭代来获得当前接收信号块更精确的信道估值（一般迭代2~3次就可以获得很好的信道估值）；

6) 最后利用步骤5)所获得的更精确信道估值 $\hat{\mathbf{h}}_{1,k}$ 和 $\hat{\mathbf{h}}_{2,k}$ 来获取当前第 k 接收OFDM信息数据符号的原始比特估值。图7给出了该算法的简单流程示意图。

这里提出的信道估计和信号检测的单块迭代操作可以保证当前接收处理单元各发射接收天线之间的信道估值快速收敛，为系统数据稳定可靠传输奠定了坚实的基础，此外时变信道环境下随着最大多普勒频率偏移的增大，增加单块迭代次数虽然可以获得一定的好处，但不能无限制地提高信道估计性能（为了降低系统处理的复杂度，这里选择迭代2~3次）；而多块递归操作则可以利用当前处理单元内部的实时信道估计和以前各相关接收处理单元的信道估值历史信息来实现系统更优良的信道估计与跟踪，为各系统结构高数据率、低误码性能以及具有健壮的鲁棒特性传输提供了强大的保障，但需要注意的是信道变化越缓慢，用于递归求平均历史信道估值数目可以越多，反之信道变化越快，用于递归求平均历史信道估值数目应该越少。

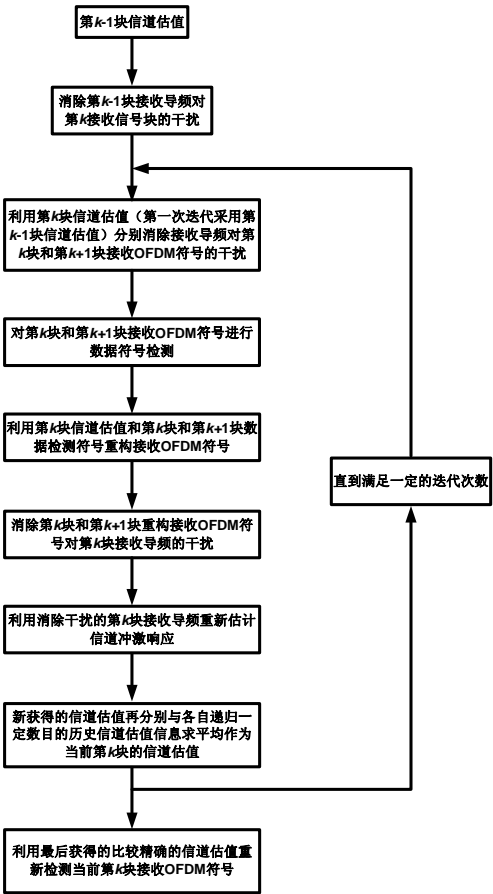


图7 时变信道估计和信号检测联合处理算法流程图

4 仿真结果和性能分析

仿真参数如下：所构建的导频辅助的无循环前缀 SFBC-OFDM 系统均采用 2×2 天线结构；各发射天线和接收天线之间的频率选择性衰落信道 **h** 采用典型城区 COST207 TU（表 1）的信道参数作为时变多径信道的开始主信道增益以及主信道时延；OFDM 信息数据块包含 2048 个数据符号，每个数据符号均来自于对信源比特流进行的 1/2 码率、约束长度 7、[171 133]的卷积编码，随机交织和 16QAM 调制；发射端两根发射天线发射的导频序列分别采用互为 127 位循环移位且总长 255 的 *m* 序列^[11]，信道估计长 127 个信号采样点；基带采样率 7.56MHz（系统带宽 8MHz，留出 5%的滚降余量）；系统接收端符号定时和载波频偏完美估计，不考虑估计误差。

表 1 COST207 典型城市信道模型

路径	时延(us)	功率增益(dB)
1	0	-3
2	0.2	0
3	0.5	-5
4	1.6	-6
5	2.3	-8
6	5	-10

图 8 给出了 non-CP SFBC-OFDM 系统在静态多径信道下、最大多普勒频率偏移为 30Hz 和 100Hz 时变信道环境下的系统误码率性能仿真对比曲线，其中各图中的“*fd*”表示最大多普勒频率偏移，它取决于终端的移动速度与载波频率，以国内移动电视广播最常用的 800MHz 频段为例，130km/h 移动速度（对应高速公路车辆行驶速度）多普勒频率偏移大约 100Hz，

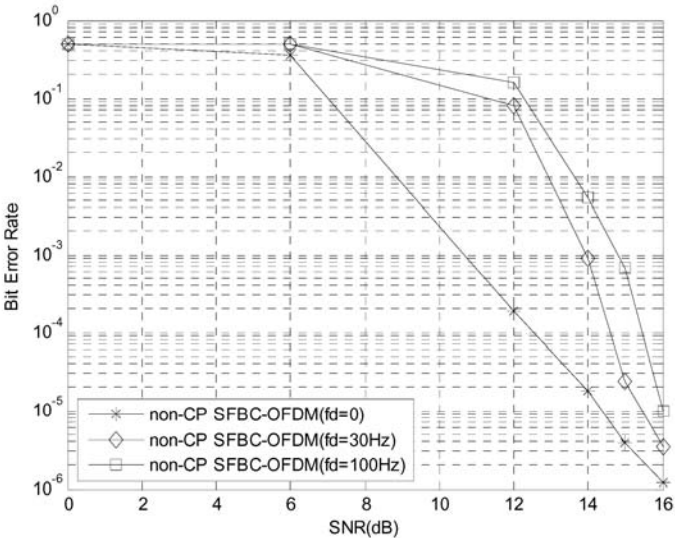


图 8 时变信道下存在不同最大多普勒频移 non-CP SFBC-OFDM 系统性能对比

40km/h 移动速度（对应市区车辆行驶速度）多普勒频率偏移大约 30Hz。从各系统仿真结果的整体趋势可以看出,随着最大多普勒频率偏移的增大(从静态信道到最大多普勒频率偏移 30Hz 和 100Hz),各系统误码率仿真性能依次下降,这说明随着时变信道冲激响应在接收信号块内变化越来越快,系统跟踪信道的性能有所下降,影响了系统的信号检测性能,但随着系统信噪比的增大,带有不同多普勒频率偏移的各系统结构还是能够满足新一代宽带无线通信系统在快速移动环境下的数据传输需求的。此外从该系统误码率性能仿真结果还可以看出:由于低信噪比下各发射接收天线信道估值没有进行任何信道估计后处理操作,因此信道噪声极大地干扰了信道估计性能,从而对信道变化的跟踪性能较差,导致各系统误码性能缓慢下降;但在中高信噪比下,信道噪声对系统的信道估计和信号检测影响变得越来越小,并且没有进行估计后处理的各发射接收天线之间的信道估值不会丢失信道变化过程中多径信道冲激响应的任何有用信息,因此各系统利用当前导频序列所获得的信道估值可以很好地估计和跟踪信道变化,使得各系统误码性能得到迅速改善,可以很好地满足终端用户对于快速移动下要求无线高速宽带数据传输的性能需求。

图 9 对比了 non-CP STBC/SFBC-OFDM 系统在静态多径信道下、最大多普勒频率偏移为 30Hz 和 100Hz 时变信道环境下的系统误码率性能,从中可以发现:在 $f_d=0$ 即静态多径信道环境下,由于 SFBC 编码的单载波或多载波系统假定在相邻频域子载波频点上多径信道频域响应近似相等存在一些误差^[12],使得其系统性能在此环境下略差于 STBC 编码的单载波或多载波系统;然而随着最大多普勒频率偏移的增大,要求在前后两个信号块时隙内保持信道基本特性不变的 STBC 编码的单载波或多载波系统越来越难以跟踪信道的实时变化,信道估计误差变得越来越大,相反不需要前后信号块联合处理的 SFBC 编码的单载波或多载波系统虽然相邻频点上的信道频域响应存在一些误差,但相比快速变化信道来说对系统性能的影响已经不明显了,表现出对快速变化的多径信道很强的适应性,从而获得了相对更好的系统误码

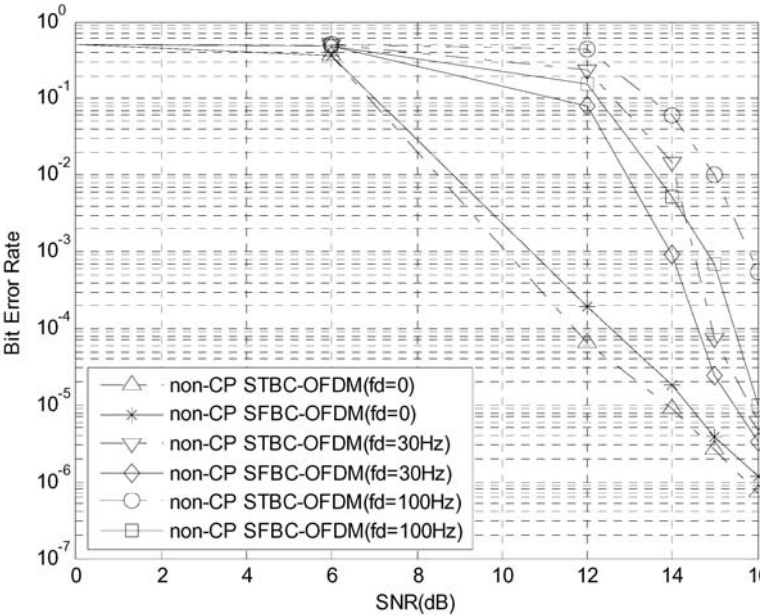


图 9 时变信道下 non-CP STBC/SFBC-OFDM 系统性能对比

率性能。因此,综合来说在时变多径信道环境下,采用本文提出的这种信道估计和信号检测联合处理算法的无循环前缀 SFBC-OFDM 表现出了更优良的系统性能,可以很好地实现无线宽带通信数据的高速率移动接收。

5 结论

本文针对时变多径信道环境下基于导频辅助的无循环前缀 SFBC-OFDM 系统,提出了一种导频和数据相互影响的干扰抑制技术即信道估计和信号检测的单块迭代和多块递归干扰抑制算法,解决了由于导频和数据相互干扰以及信道快速时变而造成该系统性能大大降低甚至无法正常工作的问题,从而为该系统在未来无线移动通信系统中广泛实际应用提供了可靠的保障,为无线高速通信数据稳定传输奠定了坚实的基础。

参 考 文 献

- [1] Stuber G. L., Barry J. R., and McLaughlin S. W., "Broadband MIMO-OFDM Wireless Communications[C]," Proceedings of the IEEE, Volume 92, Issue 2, Feb. 2004, Page(s):271-294.
- [2] Yang Hongwei, "A road to future broadband wireless access: MIMO-OFDM-Based air interface[J]," IEEE Communications Magazine, Volume 43, Issue 1, Jan. 2005, Page(s):53-60.
- [3] Kim Jaekwon, Heath R.W. Jr., Powers E.J., "Receiver designs for Alamouti coded OFDM systems in fast fading channels[J]," IEEE Transactions on Wireless Communications, Volume 4, Issue 2, March 2005, Page(s):550-559.
- [4] Shao Lei, Roy S., "Rate-one space-frequency block codes with maximum diversity for MIMO-OFDM[J]," IEEE Transactions on Wireless Communications, Volume 4, Issue 4, July 2005, Page(s):1674-1687.
- [5] Torabi M., Aissa S., Soleymani M.R., "On the BER Performance of Space-Frequency Block Coded OFDM Systems in Fading MIMO Channels[J]," IEEE Transactions on Wireless Communications, Volume 6, Issue 4, April 2007, Page(s):1366-1373.
- [6] Zhang Wei, Xia Xiang-Gen, Letaief K. Ben, "Space-Time/Frequency Coding for MIMO-OFDM in Next Generation Broadband Wireless Systems[J]," IEEE Wireless Communications, Volume 14, Issue 3, June 2007, Page(s):32-43.
- [7] Cozzo C., Hughes B.L., "Joint channel estimation and data detection in space-time communications[J]," IEEE Transactions on Communications, Volume 51, Issue 8, Aug. 2003, Page(s):1266-1270.
- [8] Han Bing, Gao Xiqi, You Xiaohu, Weckerle M., "Joint channel estimation and symbol detection for SFBC-OFDM systems via the EM algorithm[C]," 2004 IEEE International Conference on Communications, Volume 6, June 2004, Page(s):3148-3152.
- [9] Vikalo H., Hassibi B., Stoica P., "Efficient joint maximum-likelihood channel estimation and signal detection[J]," IEEE Transactions on Wireless Communications, Volume 5, Issue 7, July 2006, Page(s):1838-1845.
- [10] 任术波, 尚勇, 郭俊奇, 项海格, "基于导频辅助的无循环前缀 SFBC-OFDM 系统[J]," 通信学报, 第 30 卷第 5 期, 2009 年 5 月, pp: 68-73.

- [11] Ren Shubo, Guo Junqi, Xiang Haige, "A PN-Based Channel Estimation Algorithm in MIMO-Single Carrier Frequency-Domain Equalization System[C]," International Conference on Wireless Communications, Networking and Mobile Computing, WiCom 2007, Sept. 2007, Page(s):495-499.
- [12] Lee K.F., Williams D.B., "A space-frequency transmitter diversity technique for OFDM systems[C]," IEEE Global Telecommunications Conference, GLOBECOM '00, Volume 3, Dec. 2000, Page(s):1473-1477.

作者简介:

任术波 (1976-), 男, 河北遵化人, 北京大学博士, 主要研究方向为 MIMO、OFDM、SCFDE 及空时/空频编码技术等。

郭俊奇 (1982-), 男, 陕西西安人, 北京大学博士生, 主要研究方向为 OFDMA 技术等。

项海格 (1941-), 男, 安徽太平人, 北京大学教授、博士生导师, 主要研究方向为数字通信和信号处理, 如无线和卫星通信网络、智能天线和 MIMO、OFDM 系统、信道编码、软件无线电和片上通信系统。

基于预测的OFDMA系统中中断容量的优化

吴 波 沈 骏 陈 晨 项海格

(北京大学 信息科学技术学院, 北京 100871)

摘 要: 在 OFDMA 系统的接收端存在信道估计误差, 反馈的信道状态信息 (CSI) 存在延迟的条件下, 提出了最大化中断容量算法。首先采用自回归 (AR) 模型根据反馈预测当前的 CSI, 并根据预测的 CSI, 将问题建模为在给定中断概率下中断容量最大的优化问题。随后为降低求解该优化问题的复杂度, 提出两步次优算法, 将问题分解为子载波分配和功率分配。仿真结果表明, 在接收端存在着信道估计误差为 0 dB 时, 该次优算法仍可保证系统的中断概率不超过给定的数值, 其算法的性能远优于不采用预测的算法, 与完全 CSI 情况下相比差距在 20% 以内。

关键词: OFDMA; 资源分配; 两步次优算法; 中断容量

Outage Capacity Optimization for OFDMA systems Based on Prediction

Wu Bo Shen Jun Chen Chen Xiang Hai ge

(School of Electronics Engineering and Computer Science, Peking University, Beijing 100871)

Abstract: We propose a predictive resource allocation algorithm to maximize the outage capacity for OFDMA systems, where both channel estimation error at the receiver side and feedback delay are considered. First, we employ autoregressive (AR) model to predict the channel state information (CSI) based on the feed back CSI; with the predicted CSI, we formulate the outage maximization problem with given outage probabilities. Then, to reduce the complexity, we propose a two-step suboptimal algorithm that separates subcarrier and power allocation. Simulation results show that even when the channel estimation error is 0 dB at the receiver side, the proposed algorithm can still guarantee the outage probability under a given value, and its performance is considerably better than the algorithm without prediction; the performance gap between the proposed algorithm and the optimal algorithm in the perfect CSI case is less than 20%.

Keywords: OFDM; resource allocation; two-step suboptimal algorithm; outage capacity

1 前言

正交频分复用 (OFDM) 将宽带频率选择性信道分为若干窄带平坦衰落信道, 具有抗多

径效应和脉冲噪声的性能，因此在无线通信中得到了广泛的应用^[1]。基于 OFDM 技术的正交频分多址接入（OFDMA），将 OFDM 和频分多址相结合，除具有 OFDM 的优点外，也具有灵活的资源分配和调度的性能。以往的 OFDMA 资源分配方案^{[2][3]}多假设信号的发射端和接收端都具有完全的信道状态信息（CSI），发射端可以根据接收端反馈的前向链路的 CSI 进行子载波和功率分配。然而，在实际的无线通信环境中，由于接收端存在着信道估计的误差，以及反馈延迟、反馈错误、反馈的量化误差等因素，发射端只能得到部分的 CSI。如果发射端进行资源分配不考虑 CSI 中的错误，就会造成 OFDMA 系统的性能降低，因此针对部分 CSI 反馈的 OFDMA 系统的优化设计是当前无线资源分配的一个重要的问题。

近年，文献[4]研究了信道存在量化误差的 OFDMA 系统的 CSI 反馈策略，包括随机反馈和顺序反馈策略。文献[5]推导了在等功率分配的条件下，反馈具有量化误差和延迟下的 OFDMA 吞吐量的解析表达式。文献[6]研究了在反馈延迟下，OFDMA 吞吐量最大的资源分配策略。然而，基于 OFDMA 吞吐量性能优化的文献[4-6]研究的是系统的遍历容量。遍历容量是指系统长时间平均的容量^[6]。假定信道具有遍历特性，那么容量可以通过对足够长的码字进行计算而获得。然而时延要求高的业务不适合用长码字传输，因此时延约束下系统容量多用中断概率和中断容量作为系统性能指标^[7]。中断概率指信道的瞬时容量低于信息速率的概率，中断容量是指达到给定中断概率的最大的信息速率。现有文献并没有针对 OFDMA 系统中中断容量的优化进行研究。

本文提出了在接收端存在信道估计误差和反馈信道存在延迟的条件下，OFDMA 系统中中断容量最大的资源分配方案。该方案采用信道预测技术减弱反馈延迟和估计误差的影响，并能根据预测的 CSI 和预测的均方误差，在给定用户中断概率的条件下，优化系统吞吐量。首先，本文采用自回归（AR）模型根据反馈预测 CSI。随后，本文将中断容量的优化问题建模为混合整数规划问题。为降低求解该优化问题的复杂度，本文提出次优算法，将问题分解为两步，分别进行子载波分配和功率分配。在有 K 个用户 N 个子载波的 OFDMA 系统中，该次优算法的复杂度为 $O(KN)$ 。仿真结果显示，该次优算法在接收端信道估计误差达到 0 dB 时，系统吞吐量仍可以达到完全 CSI 情况下的 20% 以内，并能够保证系统的中断概率不超过指定的数值，性能也远优于不采用预测的资源分配算法。

符号提示： \mathbf{A}^H ， \mathbf{A}^T 分别表示矩阵 \mathbf{A} 的共轭转置和转置， $E[x]$ 表示随机变量 x 的期望， \mathbf{I}_N 表示 $N \times N$ 的单位矩阵。

2 系统模型

考虑一个有 K 个用户， N 个子载波的 OFDMA 系统。假定信道增益在 η 个 OFDM 时间周期内保持不变，用户每 η 个 OFDM 时间间隔估计一次 CSI，并将估计的 CSI 反馈给基站。假定反馈信道是无误码，但具有延迟的信道。基站根据反馈的 CSI，预测出用户当前的 CSI，并根据预测的 CSI，为下行的业务传输分配子载波和功率。

2.1 信道预测

假定从基站到每一个用户的下行信道为多径衰落信道。令 T_s 为 OFDM 的符号周期，用户 k 在子载波 n 上第 $m\eta T_s$ 时刻的信道增益的基带形式可表示为

$$H_{k,n}(m) = \sum_{l=1}^{L_k} a_{k,l}(m) e^{-j2\pi\tau_{k,l}(n-\frac{N+1}{2})\Delta f} \quad (1)$$

其中 $a_{k,l}$ 和 $\tau_{k,l}$ 分别为用户 k 的第 l 径下行信道的增益和延迟, L_k 为径的个数, Δf 为相邻两个子载波的频率间隔, 满足 $\Delta f = 1/T_s$ 。信道每径的增益 $a_{k,l}$ 可建模为相互独立的零均值平稳复高斯过程, 时间上相关性满足Jakes模型^[10]

$$\phi_k(h) = \frac{E[a_{k,l}(m+h)a_{k,l}^*(m)]}{E[|a_{k,l}|^2]} = J_0(2\pi f_{k,d}\eta T_s h) \quad (2)$$

其中 $J_0(x)$ 是第一类零阶贝塞尔函数, $f_{k,d}$ 为用户 k 的多普勒频移, 即 $f_{k,d} = v_k f_0 / c_0$, 其中 v_k 是用户 k 的移动速度, f_0 是载波频率, c_0 是真空中光速。由式(1), 用户 k 的信道增益矢量 $\mathbf{H}_k = (H_{k,1}, \dots, H_{k,N})^T$ 也是零均值平稳的复高斯矢量随机过程, 自相关矩阵 $\Sigma_{\mathbf{H}_k}$ 满足

$$[\Sigma_{\mathbf{H}_k}]_{n_1, n_2} = E[H_{k,n_1} H_{k,n_2}^*] = \sum_{l=1}^{L_k} E[|a_{k,l}|^2] e^{-j2\pi\tau_{k,l}(n_1 - n_2)\Delta f}$$

令 $\mathbf{Z}_k = (Z_{k,1}, \dots, Z_{k,N})^T$ 为用户 k 估计的信道增益矢量, 假定估计的信道增益 \mathbf{Z}_k 满足 $\mathbf{Z}_k(m) = \mathbf{H}_k(m) + \mathbf{\varepsilon}_k(m)$, $\mathbf{\varepsilon}_k \sim \text{CN}(\mathbf{0}, \sigma_\varepsilon^2 \mathbf{I}_N)$, 其中 $\mathbf{\varepsilon}_k$ 为估计误差。令用户 k 的信道增益的反馈延迟为 D_k , 基站采用 Q 阶自回归(AR)模型对信道增益进行预测

$$\hat{\mathbf{H}}_k(m + D_k) = \mathbf{A}_k \mathbf{\Psi}_k(m)$$

其中 $\hat{\mathbf{H}}_k = (\hat{H}_{k,1}, \dots, \hat{H}_{k,N})^T$ 为预测的信道增益矢量, $\mathbf{\Psi}_k(m) = (\mathbf{Z}_k(m), \mathbf{Z}_k(m-1), \dots, \mathbf{Z}_k(m-Q+1))^T$ 是基站收到的用户 k 前 Q 次反馈信息, \mathbf{A}_k 为预测系数, 满足最小均方误差(MMSE)准则

$$\mathbf{A}_k = \arg \min_{\mathbf{A}_k} \sum_{n=1}^N E[|H_{k,n}(m + D_k) - \hat{H}_{k,n}(m + D_k)|^2]$$

\mathbf{A}_k 可以按照Yule-Walker公式求解^[11]

$$\begin{aligned} \mathbf{A}_k &= E[\mathbf{H}_k(m + D_k) \mathbf{\Psi}_k^H(m)] (E[\mathbf{\Psi}_k(m) \mathbf{\Psi}_k^H(m)])^{-1} \\ &= (\mathbf{\Phi}_k \otimes \Sigma_{\mathbf{H}_k}) (\mathbf{\Gamma}_k \otimes \Sigma_{\mathbf{H}_k} + \sigma_\varepsilon^2 \mathbf{I}_{QN})^{-1} \end{aligned}$$

其中 $\mathbf{\Gamma}_k$ 为 $N \times N$ 的矩阵, $[\mathbf{\Gamma}_k]_{i,j} = \phi_k(i-j)$, $\mathbf{\Phi}_k = (\phi_k(D_k), \phi_k(D_k+1), \dots, \phi_k(D_k+Q-1))$, \otimes 为Kronecker乘积。可求出预测误差 $\mathbf{v}_k(m) = \mathbf{H}_k(m) - \hat{\mathbf{H}}_k(m)$

$$E[\mathbf{v}_k] = \mathbf{0}$$

$$\Sigma_{\mathbf{v}_k} = E[\mathbf{v}_k \mathbf{v}_k^*] = \Sigma_{\mathbf{H}_k} - (\mathbf{\Phi}_k \otimes \Sigma_{\mathbf{H}_k}) (\mathbf{\Gamma}_k \otimes \Sigma_{\mathbf{H}_k} + \sigma_\varepsilon^2 \mathbf{I}_{QN})^{-1} (\mathbf{\Phi}_k \otimes \Sigma_{\mathbf{H}_k})^H$$

由于 \mathbf{H}_k 和 $\hat{\mathbf{H}}_k$ 均是零均值的复高斯分布, 因此 \mathbf{v}_k 也是零均值的复高斯分布, 预测的信道增益对实际的信道增益的条件概率也服从复高斯分布, 即 $\mathbf{H}_k | \hat{\mathbf{H}}_k \sim \text{CN}(\hat{\mathbf{H}}_k, \Sigma_k)$ 。这样, 实际的信道能量增益 $\alpha_{k,n} = |H_{k,n}|^2$ 对预测的增益 $\hat{\alpha}_{k,n} = |\hat{H}_{k,n}|^2$ 服从自由度为2的非中心卡方分布

$$f(\alpha_{k,n} | \hat{\alpha}_{k,n}) = \frac{1}{\sigma_{k,n}^2} e^{-\frac{\alpha_{k,n} + \hat{\alpha}_{k,n}}{\sigma_{k,n}^2}} I_0\left(\frac{2}{\sigma_{k,n}^2} \sqrt{\alpha_{k,n} \hat{\alpha}_{k,n}}\right) \quad (3)$$

其中 $I_0(x)$ 为第一类零阶修正贝塞尔函数, $\sigma_{k,n}^2$ 为 Σ_k 的第 n 个对角元素。

2.2 问题描述

给定预测的信道增益, 用户 k 在子载波上 n 的中断概率为

$$\xi_{k,n} = \Pr \left(\log_2 \left(1 + \alpha_{k,n} \frac{p_n}{\sigma_r^2} \right) < R_{k,n} \mid \bar{\alpha}_{k,n} \right) \quad (4)$$

其中 p_n 为子载波 n 上的发送功率, σ_r^2 为噪声功率, $R_{k,n}$ 为用户 k 在子载波 n 上所能达到的速率, 即中断容量。式(4)可改写为

$$R_{k,n}(p_n, \bar{\alpha}_{k,n}) = \log_2 \left(1 + \frac{p_n \bar{\alpha}_{k,n}(\xi_{k,n}, \bar{\alpha}_{k,n})}{\sigma_r^2} \right) \quad (5)$$

其中 $\bar{\alpha}_{k,n}$ 由中断概率 $\xi_{k,n}$ 和预测的信道增益 $\hat{\alpha}_{k,n}$ 共同决定, 满足

$$\xi_{k,n} = \Pr(\alpha_{k,n} < \bar{\alpha}_{k,n} \mid \hat{\alpha}_{k,n}) \quad (6)$$

由式(5), $\bar{\alpha}_{k,n}(\xi_{k,n}, \hat{\alpha}_{k,n})$ 可以看作是基站根据预测的CSI和中断概率得到的等效信道能量增益。由式(6), $\bar{\alpha}_{k,n}$ 关于 $\xi_{k,n}$ 和 $\hat{\alpha}_{k,n}$ 单调递增, 即, 固定所预测的信道能量增益 $\hat{\alpha}_{k,n}$, 当用户 k 在子载波 n 上所要求的中断概率增加, 基站就可以认为用户 k 可以在子载波 n 上有更高的接收信息速率, 即有更好的等效信道能量增益; 同理, 固定用户的中断概率 $\xi_{k,n}$, 当 $\hat{\alpha}_{k,n}$ 增加时, 基站同样认为用户 k 可以在子载波 n 上有更高的等效信道能量增益。

令 $\rho_{k,n}$ 为子载波分配的指示变量: $\rho_{k,n}=1$ 表示子载波 n 分配给用户 k , 反之为 0。基于中断容量最优的问题可以建模为

$$\max_{\rho_{k,n}, p_n} \sum_{k=1}^K w_k \sum_{n=1}^N \rho_{k,n} R_{k,n}(p_n, \bar{\alpha}_{k,n}) \quad (7)$$

约束条件为

$$\sum_{n=1}^N p_n \leq p_T, p_n \geq 0 \quad (8)$$

$$\sum_{k=1}^K \rho_{k,n} = 1, \forall n, \rho_{k,n} \in \{0, 1\} \quad (9)$$

其中约束条件(8)要求基站发送数据的总功率小于 p_T , 约束条件(9)限制每一个子载波 n 仅仅分配给一个用户, 系数 w_k 为常数, 表示用户 k 的优先级权重, 当 $w_k=1$ 时, 问题变为求解用户速率和最大的优化问题。

3 子载波和功率分配算法

式(7)~(9)是混合整数规划问题, 求解最优解需要指数级的复杂度, 为降低式(7)~(9)的复杂度, 本节给出次优算法。该次优算法将式(7)~(9)化简为两步: 第一步, 假定所有子载波的发送功率相同, 为每一个子载波分配合适的用户; 第二步根据第一步子载波分配结果, 进一步为每一个子载波分配合适的功率, 并确定发送速率。

3.1 子载波分配

令子载波上的发送功率满足 $p_n=p_T/N$ ，式(7)~(9)可以分解为 N 个独立的优化问题，每一个优化问题对应一个子载波上的用户分配，对于第 n 个子载波上的优化问题为

$$\max_{\rho_{k,n}} \sum_{k=1}^K \rho_{k,n} w_k R_{k,n}(p_T / N, \bar{\alpha}_{k,n}) \quad (10)$$

约束条件

$$\sum_{k=1}^K \rho_{k,n} = 1, \rho_{k,n} \in \{0,1\} \quad (11)$$

可知问题(10)-(11)的最优解为

$$\rho_{k,n} = \begin{cases} 1, & k = \arg \max_k w_k R_{k,n}(p_T / N, \bar{\alpha}_{k,n}) \\ 0, & \text{其它} \end{cases}$$

即每一个子载波选择加权速率最大的用户。当用户的权重相同，即 $w_k=1$ ，子载波分配可以化简为

$$\rho_{k,n} = \begin{cases} 1, & k = \arg \max_k \bar{\alpha}_{k,n}(\xi_{k,n}, \bar{\alpha}_{k,n}) \\ 0, & \text{其它} \end{cases}$$

即基站将子载波分配给有最高的等效信道能量增益的用户。

3.2 功率分配

令每一个子载波 n 分配的用户为 k_n ，即 $k_n=\arg \max_k \rho_{k,n}$ ，式(7)~(9)化为

$$\max_{p_n} \sum_{n=1}^N w_{k_n} R_{k_n,n}(p_n, \bar{\alpha}_{k_n,n}) \quad (12)$$

约束条件为

$$\sum_{n=1}^N p_n \leq p_T, p_n \geq 0 \quad (13)$$

根据式(5)， $R_{k,n}$ 是 p_n 的上凸函数，因此式(12)的目标函数是关于 p_n 的上凸函数。此外，约束式(13)是关于 p_n 的线性约束，因此约束条件也是一个凸集。式(12)~(13)是关于 p_n 的凸优化问题，可以采用拉格朗日乘子法求得最优解。式(12)~(13)的拉格朗日函数为

$$\begin{aligned} L &= \sum_{n=1}^N w_{k_n} R_{k_n,n}(p_n, \bar{\alpha}_{k_n,n}) - \lambda \left(\sum_{n=1}^N p_n - p_T \right) \\ &= \lambda p_T + \sum_{n=1}^N (w_{k_n} R_{k_n,n}(p_n, \bar{\alpha}_{k_n,n}) - \lambda p_n) \end{aligned}$$

其中 λ 为拉格朗日乘子。求拉格朗日函数对 p_n 的偏导为0，得

$$\frac{\partial L}{\partial p_n} = \frac{w_{k_n} \bar{\alpha}_{k_n,n}(\xi_{k_n,n}, \bar{\alpha}_{k_n,n})}{\ln 2 \sigma_r^2 + p_n \bar{\alpha}_{k_n,n}(\xi_{k_n,n}, \bar{\alpha}_{k_n,n})} - \lambda = 0$$

即

$$p_n = \frac{w_{k_n}}{\lambda \ln 2} - \frac{\sigma_r^2}{\alpha_{k_n,n}(\xi_{k_n,n}, \alpha_{k_n,n})}$$

由于 $p_n \geq 0$ ，即最优的功率分配为

$$p_n = \left(\frac{w_{k_n}}{\lambda \ln 2} - \frac{\sigma_r^2}{\alpha_{k_n,n}(\xi_{k_n,n}, \alpha_{k_n,n})} \right)^+ \tag{14}$$

其中 $(x)^+ = \max\{x, 0\}$ ，拉格朗日乘子 λ 满足方程

$$\sum_{n=1}^N \left(\frac{w_{k_n}}{\lambda \ln 2} - \frac{\sigma_r^2}{\alpha_{k_n,n}(\xi_{k_n,n}, \alpha_{k_n,n})} \right)^+ = p_T \tag{15}$$

完成功率分配后，基站子载波 n 上按照速率 $R_{k,n}(p_n, \tilde{\alpha}_{k,n})$ 发送数据。

复杂度：子载波分配需要求解每个用户在每个子载波上的速率 $R_{k,n}$ ，复杂度为 $O(KN)$ ；功率分配时，对于每一个拉格朗日乘子 λ ，可根据式 (14) 得到每个子载波的发送功率 p_n ，复杂度为 $O(N)$ ，若求解方程式 (15) 需要进行的迭代次数为 I_λ ，则功率分配的总复杂度为 $O(N I_\lambda)$ 。由于 I_λ 可以看作是常数，该算法总的时间复杂度为 $O(KN)$ 。

4 仿真结果

本节考虑算法在瑞利时变多径信道下的性能。信道采用 COST 259 典型城区信道模型 (Typical Urban channel model) ^[10]，信道每一径自相关系数由式 (2) 给出。其他仿真参数如表 1 所示。

表 1 仿真参数

参数	数值	参数	数值
信道增益 ($E[H_{k,n} ^2]$)	1	信道估计误差 (σ_ϵ^2)	1
载波频率 (f_0)	2.6 GHz	子载波间隔 (Δf)	45 kHz
信道估计时间间隔 (η)	4	预测阶数 (Q)	5
移动速度 (v_k)	30 km/h	权重 (w_k)	1
用户数 (K)	8	子载波数 (N)	64
中断概率 ($\xi_{k,n}$)	0.1		

仿真中假定了接收端有高的信道估计误差，即 $\sigma_\epsilon^2 = E[|H_{k,n}|^2]$ 。同时假定用户的反馈延迟相同，即 $D_k = D$ ，分别仿真了 $D=1, 8, 15$ 时，本文算法的系统性能。同时我们还比较了在完全 CSI 下系统吞吐量最优的算法 0，以及基站在不采用信道预测，直接根据最新接收的 CSI 进行资源分配的算法性能。其中后两者的算法都是根据反馈的 CSI，为每一个子载波选择信道增益最大的用户，再采用“注水原理”为每一个子载波分配发送功率^[11]。

图 1 给出了在不同的信噪比下，基站成功发送给用户的数据速率。该速率定义为系统的中断容量和系统处于非中断状态概率的乘积。在图中，随着反馈延迟的增加，系统所能够达到的数据速率逐渐减小。随着信噪比的增大，在存在反馈延迟下的系统速率与完全 CSI 的速率差距也会随之增加，但是 $D=1$ 下的速率与完全 CSI 的速率相比，差距仍然是在 20% 以内。

此外，如果基站直接按照反馈的 CSI 进行资源分配，在信噪比为 50 dB 时，仅为 $D=15$ 时采用本文算法数据速率的 1/9。

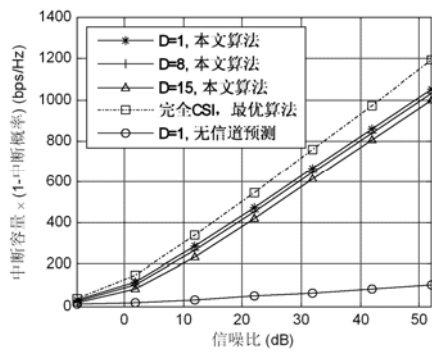


图 1 不同算法吞吐量的比较

图 2 给出了在采用不同算法时，系统实际达到的中断概率。采用本文的算法，系统的中断概率能够保持在所要求的 $\xi_{k,n}=0.1$ 附近。如果直接根据反馈的 CSI 进行子载波和功率分配，在信噪比大于 10 dB 中断概率可超过 0.9，系统的性能会恶化到几乎不可用。从图 1 和图 2 可以看出，本文的算法在接收端存在着高的信道估计误差下仍然有较好的性能，随着反馈延迟的增加，系统会降低分配给各个子载波的速率，从而保证系统的中断概率能够不超过给定的数值。

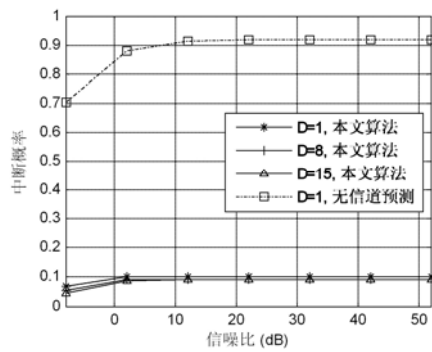


图 2 不同算法中断概率的比较

5 结论

本文提出了在接收端存在信道估计误差，反馈信道存在延迟条件下，OFDMA 系统中中断容量最大的资源分配算法。本文采用 AR 模型预测信道增益，并根据预测的 CSI 将中断容量最大的问题建模为混合整数规划问题。为降低求解复杂度，本文提出了次优算法，将问题分解为子载波分配和功率分配两个子问题，其中每一个子问题都可以求解出最优解。该次优算法总复杂度为 $O(KN)$ 。仿真结果给出，该次优算法在接收端有高的信道估计误差时，仍然有较好的性能，并能够保证系统的中断概率不超过指定的数值。

参 考 文 献

- [1] Bahai A R, Saltzberg B R, Ergen M. Multi-carrier digital communications: theory and applications of OFDM[M]. New York: Springer Verlag, 2004.
- [2] Inhyoung K, In-Soon P, Lee Y H. Use of linear programming for dynamic subcarrier and bit allocation in multiuser OFDM[J]. IEEE Transactions on Vehicular Technology, 2006, 55(4): 1195~1207.
- [3] Feiten A, Mathar R, Reyer M. Rate and power allocation for multiuser OFDM: an effective heuristic verified by branch-and-bound[J]. IEEE Transactions on Wireless Communications, 2008, 7(1): 60~64.
- [4] Agarwal R, Majjigi V, Zhu H, et al. Low complexity resource allocation with opportunistic feedback over downlink OFDMA networks[J]. IEEE Journal on Selected Areas in Communications, 2008, 26(8): 1462~1472.
- [5] Kuehne A, Klein A. Adaptive subcarrier allocation with imperfect channel knowledge versus diversity techniques in a multi-user OFDM-system[A]. IEEE PIMRC. Athens, Greece, 2007: 1~5.
- [6] Wong I C, Evans B. Optimal resource allocation in the OFDMA downlink with imperfect channel knowledge[J]. IEEE Transactions on Communications, 2009, 57(1): 232~241.
- [7] Bettesh I, Shamai S. Outages, expected rates and delays in multiple-users fading channels[C]. Proc. Conf. on Information Science and Systems. 2000.
- [8] Jakes W C. Multipath interference[J]. Microwave mobile communications, 1974.
- [9] Tsay R S. Analysis of financial time series[M]. New Jersey: Wiley-Interscience, 2005.
- [10] 3rd Generation Partnership Project. Universal mobile telecommunications system (UMTS); deployment aspects (3GPP TR 25.943 version 7.0.0 Release 7)[S], 2007.
- [11] Jiho J, Kwang B L. Transmit power adaptation for multiuser OFDM systems[J]. IEEE Journal on Selected Areas in Communications, 2003, 21(2): 171~178.

作者简介:

吴波 (1982 年—), 男, 湖北省武汉市, 北京大学信息科学技术学院在读博士研究生, 主要研究领域为 OFDMA 系统的资源分配;

沈俊 (1979 年—), 男, 上海市, 北京大学信息科学技术学院在读博士研究生, 主要研究领域为无线通信组播和单播业务的资源分配;

陈晨 (1982 年—), 男, 北京市, 北京大学信息科学技术学院博士后, 主要研究领域为网络编码、无线通信网的资源分配;

项海格 (1941 年—), 男, 安徽省太平市, 北京大学信息科学技术学院教授、博士生导师, 主要研究领域为信息论, 无线通信中的信号处理。

一种基于预滤波的NC-OFDM信号同步参数盲估计算法

吴佩霖 许小东 戴旭初

(中国科学技术大学电子工程与信息科学系, 安徽合肥, 230027)

摘要: 在存在主用户的干扰环境中, 频谱空洞的随机非连续频谱分布特征为 NC-OFDM 信号同步带来了新的技术挑战。根据有效数据子载波上的功率差异性, 提出了一种基于预滤波的 NC-OFDM 同步参数联合盲估计算法。该算法首先采用低阶滤波器, 有效抑制主用户干扰; 然后利用均衡方式在联合估计同步参数时, 补偿滤波带来的能量损失。仿真实验验证了该算法的有效性, 并表明在较强主用户干扰下, 通过预滤波和滤波均衡处理, 该算法能获得逼近于无干扰时的估计性能。

关键词: NC-OFDM; 同步参数; 盲估计; 预滤波; 均衡

A Blind Estimation Algorithm of Synchronization Parameters for NC-OFDM Signals Based on Pre-filtering

Wu Pei lin Xu Xiao dong Dai Xu chu

(Department of Electronic Engineering and Information Science, University of Science and Technology of China, Hefei, Anhui, 230027, China)

Abstract: Given the interference from primary users, the random distribution of spectrum fragments took great challenges to the synchronization for NC-OFDM systems (Non-contiguous OFDM). A joint blind estimation algorithm was proposed based on the pre-filtering and the variance of subcarrier power (VSP). The propose method utilized low-order filter to make effective suppression of the interference by primary users. Then, equalization was performed along with the joint estimation procedure to eliminate the energy loss induced by pre-filtering. Computer simulations validate the efficiency of the presented algorithm, which shows that estimation performance with such a pre-filtering and equalization processing can approximately approach those affected by no interference.

Keywords: Non-contiguous OFDM; Synchronization Parameters; Blind Estimation; Pre-filtering; Equalization

1 引言

基于认知无线电的动态频谱接入技术是缓解频谱资源日益紧张局面, 显著提高频谱利用效率的一项新技术。非连续频谱 OFDM 传输方式 (NC-OFDM) 是基于频谱池思想的 OFDM 技术^[1], 能够灵活地实现频谱资源的组合利用, 有效适应动态频谱环境。与传统 OFDM 系统

不同，它采用伺机方式使用频谱空洞，并分配相应的子载波来传输数据，而将与主用户占用频谱相对应的子载波全部关闭。这种传输方式可以实现感知用户和主用户之间的多系统共存，是未来 4G 系统的候选方案之一^[2]。

相比于传统 OFDM 系统，同步技术仍然是 NC-OFDM 的关键技术之一。错误的符号定时会引起符号间干扰 (ISI)，载波频率偏差则会造成子载波间干扰 (ICI)。已有研究表明，传统方法大多不适用于 NC-OFDM 系统^[3-8]。这主要包含两个方面的因素：第一，频谱空洞的随机非连续分布特征不利于在 NC-OFDM 系统中采用训练序列来辅助实现定时与载波同步；第二，主用户对感知用户的严重干扰妨碍了传统同步方法在 NC-OFDM 系统中的有效运用。文献[5-6]在窄带干扰条件下，研究 OFDM 的同步性能，指出 OFDM 同步性能易受用户干扰影响。文献[7-8]采用自适应预滤波，从时域探讨了 NC-OFDM 系统的同步问题，表明预滤波能较好地滤除主用户干扰，一定程度上提高了同步性能。

鉴于频谱空洞的随机非连续分布本质，以及主用户干扰的客观存在性，采用非数据辅助的方式设计 NC-OFDM 信号同步算法，不仅可以避免设计复杂的训练序列，同时还可以有效提高 NC-OFDM 系统的传输效率，提高频谱利用率。本文在文献[9-10]的基础上，提出一种同步参数联合盲估计算法，该方法通过预滤波抑制主用户干扰，再利用均衡补偿滤波造成的信号能量损失，能够灵活地利用有效子载波信号来估计同步参数。

2 问题描述

2.1 信号模型

令 NC-OFDM 符号中，子载波总数为 N ，循环前缀长度为 L_c 。定义子载波数据向量 $\mathbf{d}^i = [d_0^i, d_1^i, \dots, d_{N-1}^i]^T$ ，其中 d_k^i 表示第 i 个 NC-OFDM 符号在第 k 个子载波上承载的恒模调制符号，且满足对有效子载波有 $|d_k^i| = 1$ ，而对无效或空子载波有 $d_k^i = 0$ 。则接收端信号 $y(n)$ 可以写成

$$y(n) = e^{j2\pi\epsilon n/N} x(n) + w(n) \quad (1)$$

其中， $\epsilon = \Delta f T_0 N$ 是归一化载波频偏， T_0 是采样周期，本文中只考虑小数倍频偏，即 $\epsilon \in [-0.5, 0.5]$ 。 $w(n)$ 为零均值，方差为 σ_w^2 的加性高斯白噪声。信号 $x(n)$ 满足

$$x(n) = \sum_{i=0}^{M-1} \sum_{m=-L_c}^{N-1} s_m^i h_0(n-m-i(N+L_c)) \quad (2)$$

其中， $h_0(n)$ 是发送端滤波器，传输信道，接收端滤波器的卷积形式；而 s_n^i 是第 i 个 NC-OFDM 符号块的第 n 个样本点，即

$$s_n^i = \frac{1}{\sqrt{N+L_c}} \sum_{k \in I_d} d_k^i e^{j2\pi kn/N} \quad -L_c \leq n \leq N-1 \quad (3)$$

式中， I_d 表示 NC-OFDM 的有效子载波集合。

假设接收端能获得 NC-OFDM 信号的有效定时（即定时点 θ 落在有效 CP 以内），则根据 CP 的周期性，式(3)可以改写为向量形式

$$\mathbf{y}^i = [y_{\theta}^i, \dots, y_{N+\theta-1}^i]^T = \phi^i(\epsilon) \Phi(\epsilon) \mathbf{W} \mathbf{H} \boldsymbol{\beta}(\theta) \mathbf{d}^i + \mathbf{n}^i \quad (4)$$

其中， $\boldsymbol{\beta}(\theta) = \text{diag}\{\beta_0(\theta), \dots, \beta_{N-1}(\theta)\}$ 是定时点落入有效 CP 内造成的固定相位旋转矩阵，

其中 $\beta_k(\theta) = \exp(j2\pi k\theta/N)$, $|\beta_k(\theta)|=1$, 且 θ 应满足 $\theta \in [L, L_c]$; $\phi^i(\varepsilon) = e^{j2\pi\varepsilon((i-1)N+iL_c)/N}$ 是公共的频偏相位因子; 对角矩阵 $\Phi(\varepsilon) = \text{diag}(1, e^{j2\pi\varepsilon/N}, \dots, e^{j2\pi(N-1)\varepsilon/N})$ 是载波频偏矩阵; \mathbf{W} 是 $N \times N$ 点的 IDFT 矩阵, 它的第 k 行 n 列元素可表示为 $F_{k,n} = e^{j2\pi kn/N} / \sqrt{N+L_c}$; $\mathbf{H} = \text{diag}(H_0, \dots, H_{N-1})$ 为等效信道的频率响应矩阵, 其中 $H_k = \sum_{n=0}^{N-1} h_n e^{-j2\pi kn/N}$; \mathbf{n}^i 为零均值的复高斯白噪声向量。

2.2 主用户干扰

主用户信号对 NC-OFDM 信号的干扰, 主要是由于 NC-OFDM 接收机的 DFT 操作, 相当于对主用户信号进行了“加窗”处理, 导致主用户的频谱发生能量扩散, 即 $|R_{l,k}^i|^2 \neq 0$, $k \in I_d$, 从而干扰了附近的 NC-OFDM 子载波数据^[11]。

不失一般性, 假设主用户采用单载波数字调制。依文献[6]可知, 在 NC-OFDM 接收端, 主用户信号作为干扰信号, 用 $r_l(n)$ 表示, 即

$$r_l(n) = e^{j2\pi pn/N} \sum_{m=0}^{LQ-1} b_m h_l(n-mQ) \quad (5)$$

式中, p 表示主用户在 NC-OFDM 频谱中的归一化位置; $h_l(t)$ 是主用户发射端滤波器, 传输信道及 NC-OFDM 接收端滤波器的卷积形式, 长度为 L ; b_m 是主用户的基带调制符号; Q 为主用户的过采样率。则在 NC-OFDM 接收端, 接收信号改写为

$$y(n) = e^{j2\pi\varepsilon n/N} x(n) + r_l(n) + w(n) \quad (6)$$

以向量形式表达

$$\mathbf{y}^i = [y_{\theta}^i, \dots, y_{N+\theta-1}^i]^T = \phi^i(\varepsilon) \Phi(\varepsilon) \mathbf{W} \mathbf{H} \beta(\theta) \mathbf{d}^i + \mathbf{r}_l^i + \mathbf{n}^i \quad (7)$$

其中, $\mathbf{r}_l^i = [r_l^i(0), \dots, r_l^i(N-1)]$ 为主用户干扰信号的时域形式, 其元素 $r_l^i(k)$ 表示第 i 个 NC-OFDM 符号内的第 k 个样本。因此, 接收端数据子载波的恢复, 会受到 ICI 项、噪声项以及主用户干扰项的多重影响。

针对主用户频谱能量扩散带来的干扰, 本文采用一种简单的干扰抑制方法, 即基于预滤波处理的干扰抑制方法^[2]。在 NC-OFDM 系统中, 由于收发端能准确获知主用户占用的频谱情况, 所以可以很容易地实现预滤波。但是注意到, 高阶的滤波器会导致更为严重的 ISI, 影响后续的同步算法的性能。而低阶滤波器较长的过渡带, 同样也会导致部分 NC-OFDM 子载波信号产生深衰落, 相当于损失了有效信号的能量, 因此也会在一定程度上影响同步算法的性能。

3 EVSP联合盲估计器

考虑到预滤波处理的双重影响, 我们折中采用一种简单的 3 阶 FIR 滤波器来滤除主用户的中心频谱, 并提出一种基于滤波均衡的子载波功率方差联合估计器 (Equalized VSP: EVSP), 以有效补偿滤波损耗。

利用频偏补偿因子 Φ_{μ}^* 对 \mathbf{y}^i 进行校正 (“*”表示共轭运算), 可得第 i 个 NC-OFDM 符号的频域恢复数据向量为

$$\mathbf{z}^i(\theta, \mu) = [z_0^i, \dots, z_{N-1}^i]^T = \mathbf{W}^H \Phi^*(\mu) \mathbf{y}^i = \mathbf{W}^H \Phi(\tilde{\varepsilon} = \varepsilon - \mu) \mathbf{W} \mathbf{u}^i + \mathbf{R}_l^i + \mathbf{N}^i \quad (8)$$

式中， $\mathbf{u}^i = \phi^i(\varepsilon)\mathbf{H}\beta(\theta)$ 表示仅受到固定的相位因子和信道衰落影响的有效数据； $\mathbf{N}^i = \mathbf{W}^H \Phi^*(\mu)\mathbf{n}^i$ 是频域噪声向量； $\mathbf{R}_l^i = \mathbf{W}^H \Phi^*(\mu)\mathbf{r}_l^i$ 是主用户干扰受频偏补偿后的频域形式。展开上式，则 \mathbf{z}^i 的第 k 项可表示为

$$z_k^i(\theta, \mu) = c_0 u_k^i + \sum_{\substack{l=0 \\ l \neq k}}^{N-1} c_{l-k} u_l^i = c_0 u_k^i + I_k^i + R_{l,k}^i + N_k^i(k) \quad (9)$$

其中， I_k^i 是载波间干扰 ICI 项； c_0 是第 k 个子载波数据自身的幅度衰减和相位旋转系数； c_{l-k} 是第 l 个子载波对第 k 个子载波的干扰系数， c_l 的表达式为

$$c_l = \frac{1}{N} \frac{\sin(\pi(l + \tilde{\varepsilon}))}{\sin(\pi(l + \tilde{\varepsilon})/N)} \exp\left[j\pi \frac{(N-1)(l + \tilde{\varepsilon})}{N}\right] \quad (10)$$

$N_k^i(k)$ 与 $R_{l,k}^i$ 分别为相应的噪声干扰和主用户干扰。

不考虑主用户干扰和噪声影响，则当频偏完全补偿，即 $\tilde{\varepsilon} = 0$ 时，有关系 $|z_k^i|^2 = |u_k^i|^2$ ；而当频偏没有完全补偿，由于 I_k^i 的影响， $|z_k^i|^2$ 之间会出现差异。同理，当发生定时错误（即定时点落在有效 CP 以外）时，由于受到 ISI 的影响， $|z_k^i|^2$ 之间也会呈现差异性。因此，衡量子载波功率方差，即可以获得同步参数的有效估计。这种思想，称为子载波功率方差（Variance of Subcarrier Power: VSP）。易知，这种差异性与子载波是否传输数据，即 $|d_k^i|$ 是否为 0 无关，十分适用于 NC-OFDM 信号同步。根据上分析，可以设计出一种基于 VSP 的定时与频偏联合估计器^[9-10]。

当存在主用户干扰时，有必要先使用滤波器滤除主用户的频谱，以提高同步的精度。注意到，低阶滤波器虽然能避免 ISI，但较长的过渡带同时也导致了有效信号的能量损失。因此，本文提出一种基于滤波均衡的子载波功率方差联合估计器（Equalized VSP: EVSP），以期在联合估计同步参数的同时，有效补偿滤波损耗。该估计器可构造为

$$\begin{aligned} [\hat{\theta}, \hat{\mu}] &= \arg \min_{\substack{\theta \in [0, N+L_c-1) \\ \mu \in [-0.5, 0.5)}} J_{EVSP}(\theta, \mu) \\ &= \arg \min_{\substack{\theta \in [0, N+L_c-1) \\ \mu \in [-0.5, 0.5)}} \left\{ \sum_{k \in I_d} \left[\frac{1}{M} \sum_{i=0}^{M-1} \left| \frac{z_k^i(\theta, \mu)}{H_{F,k}} \right|^4 - \left(\frac{1}{M} \sum_{i=0}^{M-1} \left| \frac{z_k^i(\theta, \mu)}{H_{F,k}} \right|^2 \right)^2 \right] \right\} \quad (10) \end{aligned}$$

式中， $H_{F,k} = \sum_{n=0}^{N-1} h_{F,n} e^{-j2\pi kn/N}$ ， $[h_{F,0}, \dots, h_{F,L_F-1}]$ 是滤波器的时域冲激响应， L_F 是滤波器的阶数。滤波均衡的优点是能较好地补偿滤波引起的有效信号的能量损失； I_d 是 NC-OFDM 传输数据的有效子载波集。EVSP 估计器由于可以灵活使用各个子载波数据，因此能方便地通过选择有效子载波来减小主用户的干扰。对上式进行 θ 域和 μ 域的二维搜索，即可获得有效定时点与频偏的联合估计。

4 仿真实验

为了衡量 EVSP 算法的有效性，本节我们采用时偏的均方误差 $MSE_{\theta} = E[|\hat{\theta} - \theta|^2]$ 来衡量定时估计的性能，并定义有效定时点的捕获概率 $P_r = P_r(-L_c + L + L_f \leq \hat{\theta} - \theta \leq 0)$ ；采用归一化频偏的均方误差 $MSE_{\varepsilon} = E[|\hat{\varepsilon} - \varepsilon|^2]$ 来衡量频偏估计的性能。

仿真条件：NC-OFDM 信号的带宽 20MHz，子载波总数 $N = 64$ ，循环前缀 $L_c = 16$ ；主

用户信号的带宽 2.5MHz。NC-OFDM 和主用户信号的调制方式均为 QPSK，分别采用滚降系数 α 为 0.25 和 0.5 的升余弦滤波器进行信号成型。不失一般性，假设主用户在 NC-OFDM 频谱的中心位置传输数据，则可以定义主用户在 NC-OFDM 频谱中占用的子载波序号集为 [29,37]；感知用户在通信时，一般会在主用户频段两侧添加适当的保护间隔，仿真中假设添加 2 个子载波的保护间隔，则 NC-OFDM 能够传输数据的有效子载波集 $I_d \in [1,26] \cup [40,64]$ 。NC-OFDM 接收端使用 3 阶的 FIR 高通滤波器来滤除主用户的中心频谱，滤波器的阻带宽度约等于主用户频谱的宽度，滤波器系数 $\mathbf{h}_f = [-0.2519, 0.4957, -0.2519]$ 。频偏 $\varepsilon = 0.24$ 。频率选择性信道满足 $E[|h_l|^2] = Ce^{-l/3}, l \in [0, \dots, L]$ ，且 $\sum_{l=0}^L E[|h_l|^2] = \sum_{l=0}^L Ce^{-l/3} = 1$ ，其中 $L = 10$ 。统计的 NC-OFDM 符号数 $M = 5$ 。仿真中信噪比 $\text{SNR} = 20\text{dB}$ ，蒙特卡罗实验次数为 2000 次。

仿真结果如图 1, 2 所示。图中的“性能界”表示无主用户干扰时的 VSP 算法的联合估计性能；当存在主用户干扰时，VSP 及 EVSP 算法均只使用有效的子载波集 I_d 。图 1 是联合估计所得时偏估计性能；图 2 是联合估计所得频偏估计性能。由图 1, 2 中可以看到，当 SIR 较低时，仅仅通过丢弃无效子载波和添加保护间隔，并不能有效地消除主用户带来的干扰，仅当 $\text{SIR} > 10\text{dB}$ 时，无滤波的 VSP 算法的联合估计性能才能较好的接近性能界。而经过预滤波消除干扰后，VSP 算法的联合估计性能接近性能界。但是，需要指出的是，低阶滤波器的使用，使得算法损失了大部分有效子载波的能量，导致 VSP 算法的估计性能与性能界仍有一定差距。而基于滤波均衡的 EVSP 算法，其联合估计性能明显优于未加均衡的 VSP 算法，且其估计精度在 $\text{SIR} > -10\text{dB}$ 时便能迅速逼近性能界。可见，经过预滤波和滤波均衡处理后的 EVSP 算法，在较低信干比下，能获得较好的同步参数联合估计性能，在较高信干比下，同步参数的联合估计精度能逼近性能界。

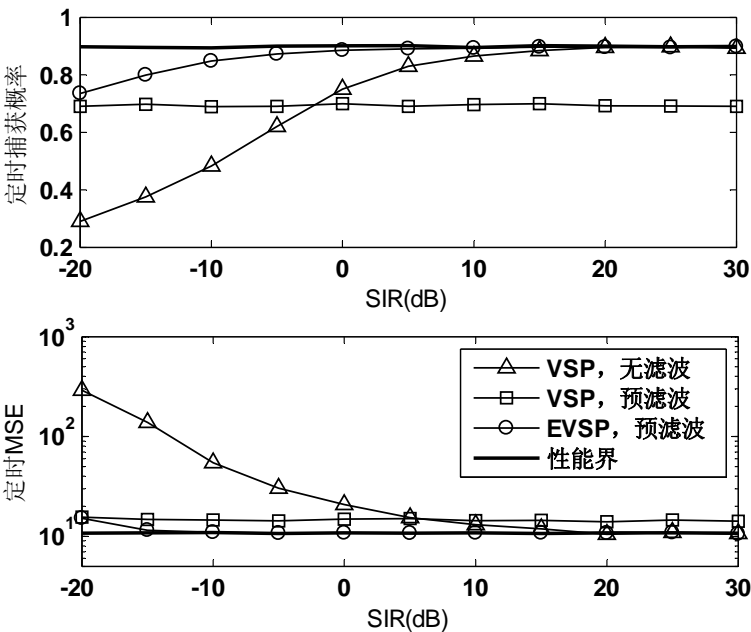


图 1 本文算法在不同信干比下的时偏估计性能

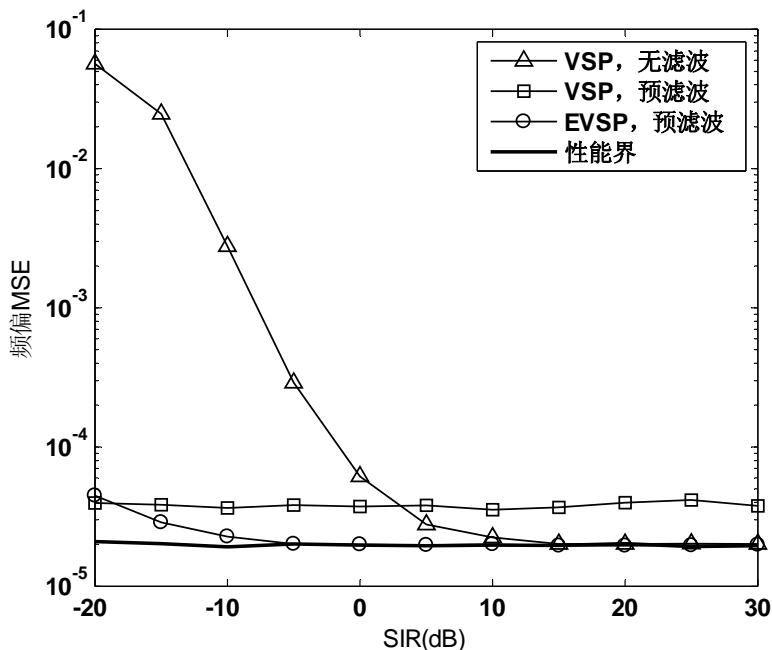


图2 本文算法在不同信干比下的频偏估计性能

5 结束语

本文针对主用户干扰环境,结合 NC-OFDM 信号频谱的随机非连续分布特点,提出一种基于预滤波的同步参数联合盲估计算法。该方法利用预滤波,对主用户干扰进行有效抑制,并采用均衡处理来补偿预滤波导致的能量损失。同时,。仿真实验表明,存在主用户干扰的通信环境中,本文提出的 EVSP 算法,能获得更稳健的联合同步性能,可获得逼近无干扰时的联合同步精度。

参 考 文 献

- [1] Poston J D, Horne W D. Discontiguous OFDM considerations for dynamic spectrum access in idle TV Channels [C]. in Proc. IEEE Int. Symp. New Frontiers Dynamic Spectr. Access Networks, Baltimore, MD, USA, 2005, 1: 607-610.
- [2] 谢显中. 感知无线电技术及其应用[M]. 北京, 电子工业出版社, 2008.
- [3] Krondorf M, Liang T J and Fettweis G. On synchronization of opportunistic radio OFDM systems [C]. 67th IEEE Vehicular Technology Conf., Marina Bay, SINGAPORE, 2008: 1686-1690.
- [4] Schmidl T, Cox D C. Robust frequency and timing synchronization for OFDM [J]. IEEE Trans. on Commun., 1997, 45(12): 1613-1621.
- [5] Marey M, Steendam H. ML frequency synchronisation for OFDM systems in the presence of overlapped narrow band transmission systems [J]. European Trans. on Telecommunications, 2009, 20(1): 97-105.
- [6] Marey M, Steendam H. Analysis of the narrowband interference effect on OFDM timing synchronization [J]. IEEE Trans. on Signal Processing, 2007, 55(9):4558-4566.

- [7] Weiss T, Krohn A, and Jondral F. Synchronization algorithms and preamble concepts in spectrum pooling systems [C]. IST Mobile and Wireless Telecommunications Summit, Aveiro, Portugal, 2003.
- [8] Acharya J, Viswanathan H and Venkatesan S. Timing acquisition for non-contiguous OFDM based dynamic spectrum access [C]. 3rd IEEE Symposium on New Frontiers in Dynamic Spectrum Access Networks, Chicago, USA, 2008: 1 – 10.
- [9] Wu L, Zhang X D, Li P S. Blind carrier frequency offset estimators based on variance for OFDM systems [C]. IEEE Int. Conf. on Communications, Glasgow, Scotland, 2007: 3024-3027.
- [10] Yao Y W, and Giannakis G B. Blind carrier frequency offset estimation in SISO, MIMO and multiuser OFDM systems [J]. IEEE Trans. on Commun., 2005, 53(1): 173-183.
- [11] Weiss T and Hillenbrand J, et al. Mutual interference in OFDM-based spectrum pooling systems [C]. 59th IEEE Vehicular Technology Conf., Milan, Italy, 2004, 4: 1873-1877.

认知无线网络频谱共享的跨层设计

张伟青 叶 梧 冯穗力

(华南理工大学电子与信息学院, 广东省广州市 510641)

摘 要: 综合研究了认知无线网络物理层、MAC 层及网络层的限制, 尤其是认知无线电特有的信道特性差异对网络拓扑的影响, 提出以最短系统激活时间为目标的跨层设计框架, 把问题公式化为混合整数线性规划(MILP)问题, 然后利用已知的信道特性计算系统所有可能的并发传送链路集合, 使 MILP 问题得以简化成线性规划(LP)问题, 从而能够以较小的计算复杂度获得全局最优解。仿真结果验证了频谱共享的跨层设计对认知无线网络系统性能的提高。

关键词: 认知无线网络; 频谱共享; 跨层设计; 异构信道

Cross-Layer Design for Spectrum Sharing in Cognitive Radio Networks

Zhang Wei qing Ye Wu Feng Sui li

(South China University of Technology, Guangzhou, 5106401, China)

Abstract: Model the channel heterogeneity, which is a unique feature in cognitive radios, and present a cross-layer design framework which jointly considers the spectrum sharing, scheduling and routing with the objective of minimizing the system total activation time, that is, maximizing the system capacity. An algorithm is proposed to find all the system concurrent transmission link-channel sets. Based on that, the original mix integer linear programming problem is then simplified to a linear programming problem, so the optimal solution could be easily found. Simulation results verify the benefit for the cross-layer design.

Keywords: Cognitive radio networks; Spectrum sharing; Cross-layer design; Heterogeneous channels

1 引言

随着无线通信的迅速发展, 高吞吐量应用的急剧增加对无线网络的容量提出越来越高的要求。另一方面, 研究表明现有的频谱分配方式(即把频谱分配给固定的授权用户或固定的业务)造成相当一部分的频率资源未能得到充分的利用^[1]。认知无线电(CR: Cognitive Radio)技术的提出使得非授权用户能够在不干扰授权用户的前提下, 感知、识别并接入当前空闲的专用频段, 从而大幅提高频谱利用率。因而, 采用认知无线电技术的认知无线网络(CRN:

Cognitive Radio Networks)有望大幅度地提高系统的容量。

对认知无线网络的研究主要集中在四个方面:频谱感知(Spectrum Sensing)、频谱管理(Spectrum Management)、频谱移动性(Spectrum Mobility)及频谱共享(Spectrum Sharing)^[2]。当前对认知无线电频谱的有效共享和分配的研究已经取得一系列的进展。如文献[3]采用了博弈论来研究频谱的共享,其研究模型是单跳多用户的场景。文献[4]考虑了多跳网络的场景,采用多信道竞争图描述各节点各信道之间相互干扰的影响,联合优化网络的频谱分配及调度,以获得公平的频谱分配。文献[5]考虑了多址接入干扰及各节点可用频带受限的影响,联合优化频谱分配、路由及调度,以最短的系统激活时间满足给定的业务需求。上述文献均未考虑由于各节点可用信道特性差异对网络设计的影响。(注:本文所讨论的信道特性差异,或称信道异构性,是指由于引入了认知无线电技术之后,系统可用频率范围变大,各不同频段的信道传播特性不同,因此可用信道带宽、覆盖范围及干扰范围也各不相同的特性)。文献[6][7]考虑了认知无线网络信道异构性的影响。文献[6]研究基于资源预留的 QOS 路由优化,为每个连接请求选择路由并分配资源,以满足不同的公平性目标。文献[7]则是联合考虑频谱共享和路由以实现最大化公平性比率的目标。文献[6][7]均把优化问题公式化为混合型整数线性规划问题(MILP: Mixed Integer Linear Programming)。MILP 通常是 NP 难(NP-Hard)问题^[8],直接求解的计算复杂度太高,只适用于小规模的网络。因此,文献[6]提出分布式路由算法,文献[7]提出启发式算法以简化问题的求解过程,但这些算法均未能获得全局最优解。

本文综合研究了认知无线网络物理层、MAC 层及网络层的限制条件,尤其是认知无线网络特有的物理层信道异构性对网络拓扑的影响,提出以最短系统激活时间满足各节点给定带宽需求为目标的跨层联合优化框架,把问题公式化为 MILP 问题,然后利用已知的信道特性计算系统所有可能的并发传送链路集合,使 MILP 问题得以简化成线性规划问题,从而能够以较小的计算复杂度获得全局最优解。仿真结果验证了我们的频谱共享的跨层设计可以明显提高认知无线网络的系统性能。

2 系统模型及问题的公式化

2.1 系统描述

本文研究的多跳认知无线网络包含了一系列固定位置的网关节点和路由器节点。一方面,来自有线网络(如 INTERNET)的业务负荷由网关节点通过无线链路分发到各路由器节点,再由各路由器节点转发给其所属的终端用户;另一方面,各路由器节点汇聚其所属各终端用户的需求,通过无线链路经由网关节点发送到有线网络。所有网关及路由器节点均配备了认知无线电功能,它们可以检测并接入当前的可用频段。当前可用的频谱资源可划分为一系列互不干扰的信道,各信道的特性(如:频率、带宽、传送/干扰范围等)各不相同。节点采用半双工模式工作,即节点在同一时间只能发送或接收数据,不能同时收发。各节点间采用单播通信,即每个传送都仅有一个发送方和一个接收方。一个节点不能同时接收来自两条链路或信道的传送。假定网络中有一个频谱服务器管理频谱的分配及调度,它周期性收集所有用户的信息(包括节点的业务需求及频段/信道的可用性信息),在每个调度周期的开始之前计算频谱分配及调度方案,并将其广播到网络中的所有节点。各节点按照接收到的方案接入相应信道。控制信息在专用的控制信道上(非认知无线电信道)交换。若服务器发现信道条件或

业务需求发生变化, 则重新计算频谱分配及调度方案。本文只讨论下行数据流(即数据从网关节点流向路由器节点)的频谱分配, 上行数据流的频谱分配可用类似方法得到。

2.2 信道模型

假设网关和路由器节点都是静止的, 信道上没有快衰落, 因此, 给定信道可支持的最大数据率(容量)是信道带宽、传输距离、发射功率及噪声功率的函数。文献[7]给出了链路 (i, j) 在信道 c 上的物理层可用容量:

$$R_{(i,j,c)} = W \log_2 \left(1 + \frac{P/d(i,j)^\eta}{N} \right) \quad (1)$$

其中, W 为信道 c 的可用带宽, P 为信道的发射功率, $d(i, j)$ 为节点 i 和 j 之间的距离, η 为路径损耗指数, N 为噪声功率。

2.3 多内容流模型

我们采用多内容流模型(Multi-commodity Flow Model)描述通过网络的数据流的多路路由, 把发往一个目的节点的业务负荷称为一个内容流(commodity), 并用其目的节点标识。

用方向图 $G = (V, E)$ 表示网络拓扑, 其中, V 为节点集合, E 为可用的链路-信道对集合, $V = (V_r + V_g)$, V_r 是网络中路由器节点的集合, V_g 是网关节点的集合。链路-信道对 $(i, j, c) \in E$ 表示节点 i 发出的数据可以通过信道 c 传送到节点 j 。用 $f_{i,j,c}^v$ 表示发往目的节点 v 的数据流在链路 (i, j) 的信道 c 上的数据流量, D_v 表示发往节点 v 的带宽需求, 则有

$$\sum_{i \in V_r, (i,v,c) \in E} f_{i,v,c}^v = D_v, \quad \forall v \in V_r \quad (2)$$

$$\sum_{i \in V_g, j \in V_r, (i,j,c) \in E} f_{i,j,c}^v = D_v, \quad \forall v \in V_r \quad (3)$$

$$\sum_{i \in V_r, j \in V_r \setminus \{v\}, (i,j,c) \in E} f_{i,j,c}^v - \sum_{k \in V_r, j \in V_r \setminus \{v\}, (j,k,c) \in E} f_{j,k,c}^v = 0, \quad \forall v \in V_r \quad (4)$$

其中, (2)式表示通过连接到节点 v 的所有链路的所有信道流入 v 的内容流的总数据量必须等于节点 v 的带宽总需求 D_v 。(3)式表示从所有源节点(网关)经过与其相连的所有链路流出的目的节点为 v 的内容流总量必须等于节点 v 的带宽总需求 D_v 。(4)式表示对于中继节点, 必须满足流守恒定律, 即流入该中继节点的关于节点 v 的内容流总量必须等于流出该中继节点的关于节点 v 内容流总量。

2.4 异构信道的传输及干扰模型

因为考虑了信道的异构性, 干扰不仅与节点的位置有关而且与节点所用信道相关, 所以我们采用链路-信道对来定义干扰。假设只要接收端无干扰, 则可传送数据。故节点 i 到 j 具有可用的链路-信道对 (i, j, c) , 当且仅当: (1)节点 i 和 j 具有相同的可用信道 c ; (2)节点 i 和 j 之间的距离不超过节点 i 在信道 c 上的发送范围; (3)所有可能干扰到节点 j 的节点 k 不发送数据。此外, 我们还定义两个链路-信道对 (i, j, c) 和 (x, y, z) 相互干扰, 若(1) $(i, j) \cap (x, y) \neq \emptyset$; 或(2) $c = z$ 且 $d(i, y) \leq I_i^c$ 或 $d(x, j) \leq I_x^z$, 其中 I_i^c 表示节点 i 在信道 c 的干扰范围。条件(1)表达了链路 (i, j) 和 (x, y) 不能有共同的收/发节点的限制, 条件(2)则包含了半双工、单播及同信道干扰的限制。若两个链路-信道对相互干扰, 则它们不能同时激活, 否则相应的传送将会失

败。我们把在任一给定时间可同时激活而互不干扰的链路—信道对集合称为传送模式 (transmission mode), 用 $\Gamma_{t,(i,j,c)} \in \Gamma$ 表示。若 (i,j,c) 在调度模式 t 中可以激活则 $\Gamma_{t,(i,j,c)} = 1$, 否则 $\Gamma_{t,(i,j,c)} = 0$ 。假定共有 N 个可能的传送模式, M 个可能的链路—信道对, 则系统可能的传送模式集合 Γ 即为 $N \times M$ 的调度矩阵。令 T_t 为分配给第 t 个传送模式的时间间隔, 则所有 N 个传送模式所需的总时间 (即系统激活时间) 为 $\sum_{t=1}^N T_t$, 有

$$\sum_{v \in V_r} f_{i,j,c}^v \leq \sum_{t=1}^N T_t \cdot \Gamma_{t,(i,j,c)} \cdot R_{(i,j,c)}, \quad \forall (i,j,c) \in E \quad (5)$$

式(5)表示所有流经链路—信道对的内容流总和不能超过该链路的实际可用容量 (链路—信道对在整个调度周期的总激活时间乘以其物理层容量), 其中物理层容量 $R_{(i,j,c)}$ 由(1)式得到。

2.5 优化问题的公式化

我们定义系统总吞吐量为所有网关节点在某个特定时间间隔 (如一个调度周期) 下发的数据总量与分配给各相关节点用于传送这些数据的总传送时间 (即系统激活时间) 的比值。显然, 给定到各路由器节点的负荷 (带宽需求), 若系统激活时间越小则系统总吞吐量就越大, 亦即系统性能就越好。所以, 我们的优化目标为在满足网络中所有节点带宽需求的前提下, 最小化系统激活时间。因此, 优化问题可公式化为:

$$\min \sum_{t=1}^N T_t \quad (6)$$

限制条件: (2)(3)(4)(5)

解出此优化公式得到的优化变量 T_t 和 $f_{i,j,c}^v$ 即可确定网络层的路由方案 (即发往目的节点的数据应该经由哪些链路传送, 各链路应传送多少的数据量)、MAC 层的调度方案 (即在某个时隙应该激活哪些链路, 各时隙最优的激活时间是多长) 及物理层的信道分配方案 (各激活的链路应采用哪些信道传送)。

3 优化问题求解

问题求解分成两步: 第一步根据节点上报的信息计算节点关联矩阵、系统所有可用链路—信道对集合及所有可能的并发传送模式集合, 构建调度矩阵; 第二步利用现有线性优化算法^[9]或优化工具 (如 MATLAB 优化工具 LINPROG) 求上节给出的线性规划问题的最优解。

1) 计算节点关联矩阵

频谱分配服务器根据各节点上报的信息计算各节点的可送达节点集和干扰节点集, 生成节点关联矩阵 I 。其中, $I_{i,j,c} \in I$ 。若节点 j 在节点 i 使用信道 c 的传送范围内, 则 $I_{i,j,c} = 1$; 若节点 j 在节点 i 使用信道 c 时的传送范围之外, 而干扰范围之内, 则 $I_{i,j,c} = 2$; 若节点 j 在节点 i 使用信道 c 时的干扰范围之外, 则 $I_{i,j,c} = 0$ 。

2) 计算系统可用链路—信道对集合

根据节点关联矩阵 I 计算所有可用链路—信道对集合 E 。其中, 若 $I_{i,j,c} = 1$, 则 $(i,j,c) \in E$ 。

3) 计算系统所有可能的并发传送集 (调度矩阵)

本算法的基本思想是从有序链路—信道对集合中选取一个元素在当前的调度时隙中激

活，并确定该元素激活时不能同时激活的所有其它链路—信道对，然后，在剩下的未确定的链路—信道对集合中再次选取一个可激活的元素，并确定受该元素影响的所有链路—信道对。循环上述过程，直到所有链路—信道对的状态（激活或不能激活）都得到确定，即得到一个时隙的调度方案。用上述方法遍历系统的链路—信道对集合所得到的调度时隙集合，即为系统所有可能的并发发送集。算法的具体实现步骤如下：

步骤 1 创建一个临时的有序链路—信道对集合 $E_1 = E$ 。

步骤 2 将 E_1 中第一个可用链路—信道对 (i, j, c) ，从 E_1 中移除，并置相应的调度矩阵元素 $\Gamma_{t,(i,j,c)} = 1$ ，即该链路在本次调度时隙 t 可以激活。对于不在剩余链路—信道对集合 E_1 中的元素 (x, y, z) ，置 $\Gamma_{t,(x,y,z)} = 0$ ，即该链路在本次调度时隙 t 中不能激活。创建一个临时的集合 E_2 ，令 $E_2 = E_1$ 。

步骤 3 对集合 E_2 中的每一元素 (x, y, z) ，若满足以下条件之一，则将该元素从 E_2 中移除，并置 $\Gamma_{t,(x,y,z)} = 0$ ：

- (1) $(i, j) \cap (x, y) \neq \emptyset$ ，即，链路 (i, j) 和 (x, y) 有共同的收/发节点；
- (2) $I_{i,y,c} \neq 0$ ，即，节点 y 在节点 i 的信道 c 的干扰范围之内；
- (3) $I_{x,j,z} \neq 0$ ，即，节点 j 在节点 x 的信道 z 的干扰范围之内；

步骤 4 若 $E_2 \neq \emptyset$ ，则置 E_2 中的第一个元素 (x, y, z) 所对应的 $\Gamma_{t,(x,y,z)} = 1$ ，置 $(i, j, c) = (x, y, z)$ ，并将该元素从 E_2 中移除。

步骤 5 若 $E_2 \neq \emptyset$ ，则转步骤 3，否则转步骤 6。

步骤 6 输出本次调度时隙 t 的调度方案 $\Gamma_{t,(x,y,z)}$ ， $(x, y, z) \in E$ 。

步骤 7 若步骤 6 输出的调度方案中有不只一个的非零元素，设第一个非零元素为 $\Gamma_{t,(i,j,c)}$ ，第二个非零元素为 $\Gamma_{t,(x,y,z)}$ ，则清空 E_2 ，然后复制 $\Gamma_{t,(x,y,z)}$ 所对应的集合 E 中元素 (x, y, z) 之后的所有元素到 E_2 ，并复制 $\Gamma_{t,(x,y,z)}$ 之前的值到新的调度方案，而置该元素值在新调度方案中所对应的值 $\Gamma_{t+1,(x,y,z)} = 0$ ，再转步骤 3 以计算并输出新调度方案。

步骤 8 若 $E_1 \neq \emptyset$ ，则返回步骤 2；否则，结束调度矩阵的计算。

4 仿真结果

我们应用上述算法研究一个拓扑随机生成的 100×100 单位长度区域的认知无线网络，其中网关节点位于区域中心，路由器节点则在区域内随机分布。我们采用 MATLAB 作为仿真工具，并用其优化工具 LINPROG 求解各线性规划方程。假定当前的网络中有两个可用频段，各频段的信道特性如表 1 所示。我们分别对不同节点数的网络进行仿真计算。其中，针对每种节点数，我们分别随机产生 200 个不同的网络。在这些网络中，各路由器节点的位置、各节点当前的可用信道及各目的节点的带宽需求均是随机产生的。

为了验证考虑频谱共享的跨层设计对网络性能的影响，我们将仿真结果与不考虑频谱共享的逐跳激活的跨层设计算法进行对比，结果如图 1 所示。我们所采用的逐跳激活算法与频谱共享算法的唯一区别在于前者不考虑频谱在相同时间不同节点间的复用，即在每个特定的时间段内系统仅激活一条链路，且在该时段该链路使用所有可用信道传送数据。我们定义归一化的系统性能提高比率为逐跳激活与频谱共享的最短系统激活时间差与频谱共享的最短系

统激活时间的比值，图 2 给出了不同节点数网络频谱共享算法对系统性能提升的变化。从图中可以看出，随着网络节点数的增加，频谱共享对网络性能的提高越来越明显。对 12 个节点的网络，采用频谱共享平均可以缩短约 10%的系统激活时间；而对 20 个节点的网络，采用频谱共享则平均可以缩短约 20%的系统激活时间。这是因为网络中可并发传送的链路随着节点数的增加而增加，所以频谱共享带来的性能提升也越来越显著。

表 1 系统可用频段及其信道参数

频段	频谱范围(MHz)	覆盖半径	干扰半径	信道容量	可用信道
A	[2400,2500]	30	60	20	1,2,3,4
B	[5700,5800]	20	40	30	5,6,7,8,9,10

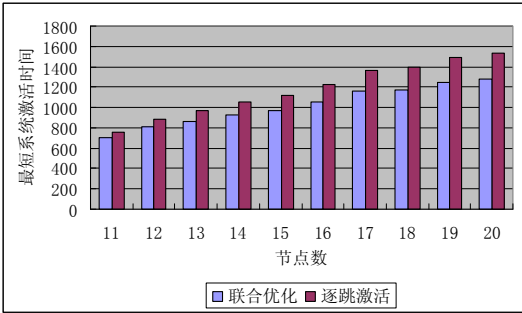


图 1 不同节点数网络频谱共享与逐跳激活方式最短系统激活时间的比较

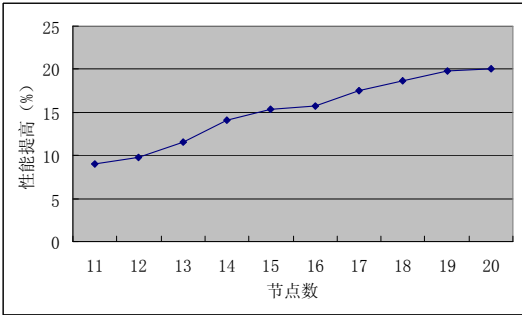


图 2 不同节点数网络归一化的频谱共享系统性能提高比率

5 结论

本文提出了一个跨层优化框架用于联合设计具有异构信道特性的认知无线网络的频谱分配、调度及路由。我们的研究解决了两个问题：（1）给定一个随机的网络及所有节点的业务需求，系统能在多短的时间内满足所有节点的业务需求？（2）为了获得最优解，如何选择激活链路，如何为这些激活链路分配频谱，如何路由各节点的业务需求？与之前的研究不同的是我们在考虑了信道异构性（带宽及传送/干扰范围的变化）对网络拓扑影响的前提下，利用已知的信道特性计算出系统所有可能的并发传送链路集，使得 MILP 的优化问题可以简化成普通的线性规划问题，从而能够以较小的计算复杂度获得全局最优解。仿真结果验证了我们提出的频谱共享的跨层设计可以在一定程度上提高认知无线网络的性能。

参 考 文 献

- [1] FCC, ET Docket No 03-222 Notice of proposed rule making and order, December 2003.
- [2] I. F. Akyildiz, W-Y Lee, M. C. Vuran and S. Mohanty, NeXt generation/dynamic spectrum access/cognitive radio wireless networks: a survey, Journal of Computer Networks, Vol. 50, No. 13, 2007, pp. 2127–2159.
- [3] CLEMENTS N, ROSE C. Intelligent power allocation strategies in an unlicensed spectrum. Baltimore: Proc IEEE Symposium on New Frontiers in Dynamic Spectrum Access Networks, 2005.
- [4] J. Tang, S. Misra and G. Xue, Joint spectrum allocation and scheduling for fair spectrum sharing in cognitive radio wireless networks, Computer Networks, Vol. 52, No. 11, 2008, pp. 2148-2158.
- [5] J Zhang, Z Zhang, H Luo, A Huang, A Column Generation Approach for Spectrum Allocation in Cognitive Wireless Mesh Network, IEEE Globecom'2008.
- [6] R. Hincapie, J. Tang, G. Xue and R. Bustamante, QoS routing in wireless mesh networks with cognitive radios, IEEE Globecom'2008.
- [7] M Ma, DHK Tsang, Joint design of spectrum sharing and routing with channel heterogeneity in cognitive radio networks, Physical Communication, 2009 Elsevier
- [8] M.R. Garey and D.S. Johnson, Computers and Intractability: A Guide to the Theory of NP-completeness, W.H. Freeman and Company, pp. 245–248, New York, NY, 1979.
- [9] M.S. Bazaraa, J.J. Jarvis, H.D. Sherali, Linear Programming and Network Flows, third ed., John Wiley & Sons, 2005.

作者简介:

张伟青 (1974-), 男, 在读博士研究生, 研究方向为宽带无线网络 Email: zhangweiqing@21cn.com

叶梧 (1938-), 男, 教授, 博士生导师, 研究方向为网络理论与技术, 数字音/视频处理, 宽带无线通信 Email: eewuye@scut.edu.cn

冯穗力 (1955-), 男, 教授, 博士生导师, 研究方向为信息网络理论与技术 Email: fengsl@scut.edu.cn

多射频多信道无线Mesh网络的跨层公平资源分配

黄鑫¹ 冯穗力¹ 庄宏成²

(1、华南理工大学 电子与信息学院, 广东 广州 510640;

2、华为技术有限公司, 广东 深圳 518129)

摘要: 在多射频多信道无线 Mesh 网络中传输业务流, 需要解决四个重要问题: 构建逻辑拓扑、给业务流确定传输路由、给链路分配信道和带宽, 因此属于一个资源分配问题。本文将跨层公平资源分配问题描述成一个整数线性最大-最小公平带宽分配问题, 并提出一种次优化的基于禁忌搜索的迭代本地查找算法来求解该问题。仿真结果表明, 所提出的资源分配机制能够保证链路间的公平性, 且相比迭代本地查找算法具有更好的收敛性和更优的性能。同时研究了网络资源对吞吐量和公平性的影响, 可以为网络资源配置提供参考。

关键词: 无线 Mesh 网络; 跨层; 资源分配; 最大最小公平性

Cross-layer Fair Resources Allocation for Multi-Radio Multi-Channel Wireless Mesh Networks

Huang Xin¹ Feng Sui li¹ Zhuang Hong cheng²

(1. School of Electronic and Information Engineering, South China University of Technology, Guangzhou 510640, China; 2. Huawei Technologies Co. Ltd., Shenzhen 518129, China)

Abstract: When transmitting traffic flows in a multi-radio multi-channel wireless mesh network, there are four important issues should be addressed: logical topology formulation, routing, channel allocation and bandwidth assignment. It's a resources allocation problem. In this paper, we mathematically formulate the cross-layer fair resource allocation problem as an integer linear max-min fair bandwidth allocate problem, and propose a near optimal Tabu Search based Iterated Local Search (TSILS) algorithm to solve the problem. Simulation results show that the proposed resources allocation scheme achieves good fairness among links, and it achieves better astringency and performance than Iterated Local Search (ILS) algorithm. Performance impact of available network resources is also assessed in terms of aggregate network through and fairness index, which gives advice for the resources deployment of the network.

Keywords: Wireless mesh networks; cross-layer; resources allocation; max-min fairness

1 引言

无线 Mesh 网络(Wireless Mesh Networks: WMNs)具有分布式、部署容易、鲁棒性好以及覆盖范围广的特点, 因此成为研究的热点。一个 WMNs 通常由静态的 Mesh 路由器和移动的

Mesh 客户端组成。Mesh 路由器以无线多跳的方式相互连接，其中一些路由器作为网关通过高速有线链路连接到核心网。每个 Mesh 路由器既能为其客户端传输业务，也能为其邻居中继业务。使用多射频和多信道能够大大增加 WMNs 的带宽^{[1][2]}。在一个多射频多信道的 WMNs (multi-radio and multi-channel WMNs: MRMC-WMNs)中，每个 Mesh 路由器配置有多个射频，网络中有多个正交频率信道可用。如果两个邻居 Mesh 路由器都有一个射频工作在相同的信道上，则它们之间可以建立一条逻辑链路相互通信。

WMNs 通常没有固定基础设施，网络参数具有动态和时变特征，因此传统的基于层的协议设计不能为其提供性能优化机制。跨层设计方法可以联合设计多个协议层来优化系统性能，使得网络能在参数动态变化和资源有限的情况下为用户提供更好的服务质量保证^{[3][4]}。Akyildiz 等人总结了跨层优化的必要性，且研究了相应的跨层机制和算法^[4]。Mung Chiang 等人总结了跨层优化分解的数学理论^[5]。文献[2]将 WMNs 的逻辑拓扑设计，射频分配、信道分配和路由建模为一个混合整数线性规划问题，优化目标为负载均衡和拥塞控制。然而，当传输负载较重时，所提出的迭代本地查找(Iterated Local Search: ILS)算法的初始迭代不可行。Mohsenian 等人将联合信道分配和拥塞控制建模为一个网络效用最大化问题^[6]，且在文献[7]中研究了联合信道分配、射频分配和媒体接入控制设计问题。但是，文献[6]和文献[7]主要考虑了增加网络吞吐量，没有考虑公平性。Alicherry 等人对联合信道分配和路由问题进行数学建模，并提出一种常数近似算法来求解该 NP 难问题，该算法能够在满足公平性的限制下最大化网络吞吐量^[8]。然而，文章没有分析所获得吞吐量的公平性，且算法非常复杂。在本文中，我们将联合逻辑拓扑设计、路由、信道分配和带宽分配所涉及到的跨层资源分配问题建模为一个整数线性最大-最小公平带宽分配问题，并且提出一个次优化的算法来获得其近似最优解。

2 系统描述和问题建模

2.1 网络模型

可以用物理连接图 $G(N, E)$ 表示一个 MRMC-WMN，其中 $N = \{1, 2, \dots, |N|\}$ 表示所有无线 Mesh 路由器的集合， E 表示所有的方向性边集合。在后文中，术语 Mesh 路由器和节点交替使用。对任意两个节点 $a, b \in N$ ，如果节点 b 在节点 a 的通信范围之内，则有一条从节点 a 到节点 b 的边 $e_{ab} \in E$ 。假设所有节点的通信范围相同，因此边的连接具有对称性。即如果有 $e_{ab} \in E$ ，则必须存在 $e_{ba} \in E$ 。对于 $\forall a \in N$ ，令 I_a 表示其配置的射频数目。网络中有 C 个正交频率信道可用，令 \mathbf{C} 表示可用的信道集，有 $\mathbf{C} = \{1, \dots, C\}$ 。MRMC-WMNs 拓扑可以由一个节点-边关联矩阵 $A \in R^{|N| \times |E|}$ 表示，定义为：

$$A(n, e) = \begin{cases} -1, & \text{节点 } n \text{ 是边 } e \text{ 的发送节点,} \\ 1, & \text{节点 } n \text{ 是边 } e \text{ 的接收节点,} \\ 0, & \text{其它.} \end{cases} \quad (1)$$

2.2 信道分配

对任意边 $e_{ab} \in E$ 和任意信道 $i \in \{1, \dots, C\}$ ，定义一个二进制的链路信道分配变量 l_{ab}^i 。如果信道 i 分配给了边 e_{ab} ，则 l_{ab}^i 等于 1；否则 l_{ab}^i 等于 0。 $l_{ab}^i = 1$ 表示节点 a 建立了一条工作在信

道 i 上的逻辑链路 (a,b) 来与节点 b 通信。根据连接的对称性, 存在 $l_{ab}^i = l_{ba}^i, \forall a,b \in N, e_{ab} \in E, \forall i=1, \dots, C$ 。我们允许在两个邻居节点之间建立多条逻辑链路通信, 这些链路工作在不同的信道上, 因此能够增大两个邻居节点之间的有效带宽。

对任意节点 $a \in N$ 和任意信道 $i \in \{1, \dots, C\}$, 定义一个二进制的节点信道分配变量 v_a^i 如下: 如果 $\exists b \in N$ 且 $e_{ab} \in E$ 使得 $l_{ab}^i = 1$, 则 $v_a^i = 1$; 否则 $v_a^i = 0$ 。 $\sum_{i=1}^C v_a^i$ 表示节点 a 用于与其邻居建立逻辑链路的信道数目。每个射频必须工作在不同的信道上, 因此存在节点-射频限制 $\sum_{i=1}^C v_a^i \leq I_a, \forall a \in N$ 。链路信道分配向量 l_{ab}^i 和节点信道分配向量 v_a^i 提供了逻辑拓扑设计和信道分配所要求的所有信息。

2.3 带宽分配

令 c_p 表示 WMNs 的物理层链路带宽; c_{ab}^i 表示逻辑链路 (a,b) 在信道 i 上的有效链路带宽, 其取决于链路的激活时间。存在 $c_{ab}^i \leq l_{ab}^i c_p, \forall a,b \in N, e_{ab} \in E, \forall i=1, \dots, C$ 。逻辑链路的激活受到两类限制: 第一, 节点不能在同一信道上同时发送和接收数据; 第二, 节点数据的发送和接收可能受到其邻居节点的干扰。当 $l_{ab}^i = 1$ 时, 节点 a 干扰范围内的所有节点都不能在信道 i 上接收数据, 且节点 b 干扰范围内的所有节点都不能在信道 i 上传输数据。定义节点 a 的干扰节点集为 $N(a)$, 链路 e_{ab} 的干扰链路集为 I_{ab} , 存在 $I_{ab} = \{e_{mn} \mid \forall m \in N(b) \text{ 或 } n \in N(a)\} - \{e_{ab}\}$ 。链路 e_{ab} 和其干扰链路不能同时激活, 因此其需要共享物理层的链路带宽。存在带宽分配限制^[9,10]: $c_{ab}^i + \sum_{e_{mn} \in I_{ab}} c_{mn}^i \leq c_p, \forall a,b \in N, e_{ab} \in E, \forall i=1, \dots, C$ 。其中 c_{ab}^i/c_p 表示逻辑链路 (a,b) 在信道 i 上激活的时间。

2.4 路由

假设网络中有 K 个流, 令 $F = \{1, 2, \dots, K\}$ 表示流的集合, 令 f_k 表示流 k 的流量。业务流能够用一个节点-流关联矩阵 $T \in R^{|N| \times |F|}$ 表示, 定义如下:

$$T(n,k) = \begin{cases} -f_k, & \text{节点 } n \text{ 是流 } k \text{ 的源节点,} \\ f_k, & \text{节点 } n \text{ 是流 } k \text{ 的目的节点,} \\ 0, & \text{其它.} \end{cases} \quad (2)$$

对于任意流 $k \in F$ 、任意链路 $e_{ab} \in E$ 和信道 $i=1, \dots, C$, 定义一个二进制的路由变量 $r_{ab,i}^k$ 。如果流 k 在信道 i 上经过链路 (a,b) 传输, 则 $r_{ab,i}^k$ 等于 1; 否则等于 0。为了保持分组到达次序和路由的简化, 本文只考虑单径路由。所以, 在任意两个邻居节点之间, 只允许用一条逻辑链路来传输一个流, 即存在 $\sum_{i=1}^C r_{ab,i}^k \leq 1, \forall a,b \in N, e_{ab} \in E, k \in F$ 。

对于任意流 k , 其传输路径长度为 $\sum_{a,b \in N, e_{ab} \in E} \sum_{i=1}^C r_{ab,i}^k$ 。令 SL_k 表示流 k 在连接图 $G(N,E)$ 中的最短路径长度。MRMC-WMNs 的公平资源分配考虑了负载均衡, 但是可能造成传输路径过长。为了避免这种情况, 定义了路径长度限制如下:

$$\sum_{a,b \in N, e_{ab} \in E} \sum_{i=1}^C r_{ab,i}^k \leq SL_k + \alpha, \forall 1 \leq k \leq K \quad (3)$$

其中, α 是一个用于限制路径长度的可调参数。通过设置 $\alpha = 0$, 路由算法总能找到最短路径。定义一个边-流关联矩阵 $R \in R^{|E| \times |F|}$, 其中 $R(e_{ab}, k) = \sum_{i=1}^C r_{ab,i}^k f_k$ 表示流 k 分配给边 e_{ab} 的业务量。因此, 存在一个路由和流分配限制: $A \square R = T$ 。令 β_{ab}^i 表示在逻辑链路 (a, b) 上通过信道 i 传输的业务总量。可知, $\beta_{ab}^i = \sum_{1 \leq k \leq K} r_{ab,i}^k f_k$, $\forall a, b \in N, e_{ab} \in E, \forall i = 1, \dots, C$, 而 β_{ab}^i 必须小于分配给逻辑链路 (a, b) 在信道 i 上的带宽 c_{ab}^i 。因此存在如下的流限制:

$$\varepsilon c_{ab}^i - \beta_{ab}^i \geq 0, \quad \forall a, b \in N, e_{ab} \in E, \forall i = 1, \dots, C, \varepsilon \leq 1 \quad (4)$$

其中 ε 是链路的最大利用率。通常, 高的链路利用率意味着高的排队延时^[10]。为了避免过载^{[11][12]}, 本文设置 $\varepsilon = 0.8$ 。可知, $\varepsilon c_{ab}^i - \beta_{ab}^i$ 表示逻辑链路 (a, b) 在信道 i 上的剩余带宽。令 ω_{\min} 表示所有逻辑链路中的最小剩余带宽, 要求:

$$\omega_{\min} \leq \varepsilon c_{ab}^i - \beta_{ab}^i + \varepsilon c_p(1 - l_{ab}^i), \quad \forall a, b \in N, e_{ab} \in E, \forall i = 1, \dots, C \quad (5)$$

2.5 优化问题和提出的算法

为了获得公平的资源分配, 本文根据最大-最小公平性限制分配带宽给存在的逻辑链路^{[13][14]}, 优化目标为最大化最小剩余带宽 ω_{\min} 。给定 MRMC-WMNs 的物理拓扑图 $G(N, E)$ 和参数 $I_a, C, c_p, I_{ab}, f_k, SL_k, \alpha, \varepsilon$, 最优化问题描述如下:

$$\begin{aligned} & \underset{l_{ab}^i, v_a^i, r_{ab,i}^k, \omega_{\min}}{\text{maximize}} && \omega_{\min} \\ & \text{subject to} && \\ & l_{ab}^i = l_{ba}^i, \quad \sum_{i=1}^C v_a^i \leq I_a \\ & c_{ab}^i \leq l_{ab}^i c_p, \quad c_{ab}^i + \sum_{m, n \in e_{mn} \in I_{ab}} c_{mn}^i \leq c_p \\ & \sum_{i=1}^C r_{ab,i}^k \leq 1, \quad \sum_{a, b \in N, e_{ab} \in E} \sum_{i=1}^C r_{ab,i}^k \leq SL_k + \alpha \\ & \varepsilon c_{ab}^i - \beta_{ab}^i \geq 0, \quad A \square R = T \\ & \omega_{\min} \leq \varepsilon c_{ab}^i - \beta_{ab}^i + \varepsilon c_p(1 - l_{ab}^i) \\ & \text{where} \\ & l_{ab}^i, v_a^i, r_{ab,i}^k \in \{0, 1\}, \quad c_{ab}^i, \omega_{\min} \geq 0 \\ & \forall a, b \in N, e_{ab} \in E, \forall i = 1, \dots, C, \forall k = 1, \dots, K \end{aligned}$$

该最优化问题是一个整数线性规划(Integer Linear Programming: ILP)问题。对于小规模 MRMC-WMNs, 该问题能够用分支定界法轻松求解^[15]。然而, 当网络规模大的时候, 求解变得非常困难。后启发式算法广泛应用于寻找近似最优解^[6]。文献[2]所提出的 ILS 算法存在两个问题: 首先, 其设置一个单信道逻辑拓扑作为初始解, 但是单信道网络的吞吐量比较低, 所以当传输负载比较重的时候, 不能满足不等式(4)的流限制, 因此初始解的本地迭代不可行; 其次, 在迭代过程中, 其放宽了二进制路由变量的整数限制, 从而每个流获得了多条路径。然而, 在多条路径中选择最大可行性的路径作为最终路由后, 算法没有重新分配带宽给逻辑链路。禁忌搜索技术被认为是一种用于优化信道分配的好工具^[16], 本文提出一种基于禁忌搜索的信道分配算法(算法 1)来获得信道分配的初始解, 其伪代码如下:

算法1：基于禁忌搜索的信道分配算法

- 1: 设置 $x=0$ 且 X =最大迭代次数.
- 2: 设置 $UC=0$ 且 UC_{\max} = 总的边数目.
- 3: 设置 $TabuLists=\emptyset$.
- 4: 设置 期望法则=优于当前最佳解
- 5: $s_{best} \leftarrow$ 对于 $\forall a, b \in N, e_{ab} \in E$ 设置 $l_{ab}^1 = 1$, 将 l_{ab}^1 更新到 $TabuLists$.
- 6: while $x \neq X$ and $UC \neq UC_{\max}$ do
- 7: 计算 $I(s_{best})$
- 8: 构建 s_{best} 的 λ 个邻近解 $\{s_1, s_2, \dots, s_\lambda\}$, 使其满足节点-射频数限制, 违反 $TabuLists$, 或满足期望法则.
- 9: $s_{temp} \leftarrow \arg \min_{s_1 \leq s_2 \leq s_\lambda} I(s_i)$.
- 10: if $I(s_{temp}) < I(s_{best})$ then $s_{best} \leftarrow s_{temp}$, $UC=0$,
- 11: else $UC=UC+1$.
- 12: 将 s_{best} 更新到 $TabuLists$.
- 13: $x=x+1$.
- 14: end

其中: s 表示信道分配的解; $I(s)$ 表示解的冲突边总数; UC 用于记录没有改善的解数目。为了获得信道分配的快速收敛, 使用禁忌链表 $TabuLists$ 来避免重新分配同样的信道给同一条边。期望法则 **Aspiration criteria** 用于允许被禁忌的好的解。邻近解是通过随机选择一条边, 并重新分配信道给该边来获得。算法 1 的目的是找出具有最小冲突边的信道分配最优解 s_{best} 。由仿真可知, 本算法能够在 50 次迭代内找到具有最小网络干扰的最优解。

基于算法 1, 本文提出一种基于禁忌搜索的迭代本地查找(Tabu Search based Iterated Local Search: TSILS)算法, 其伪代码如算法 2 所示。在第二行, 使用算法 1 获得的信道分配解决方案作为初始解。在每一次迭代, 通过放宽路由变量 r 的整数限制来简化优化问题, 设置链路信道分配变量 l 作为当前解, 选择一条具有最小剩余带宽的边 $e_{pq} \in E, p, q \in N$, 并重新分配信道给该边。简化后的优化问题具有更少的整数变量, 从而变得容易求解。本文使用[2]中的方法来给每个流确定一条最优的路由。然后, 通过求解具有固定信道分配和路由的优化问题给逻辑链路重新分配信道。由后面的仿真可知, 该算法相比 ILS 算法能够获得更加可行和优化的效果。

算法2：基于禁忌搜索的迭代本地查找算法(TSILS)

- 1: 设置迭代次数 X 。
- 2: 由算法1获得 $l_{ab}^1[1], \forall a, b \in N, e_{ab} \in E, \forall i=1, \dots, C$.
- 3: for $x=1$ to X do
- 4: 随机选择 $p, q \in N$ 使得 $e_{pq} \in E$.
- 5: 根据如下限制条件解优化问题

$$0 \leq l_{ab,i}^k \leq 1, \forall a, b \in N, e_{ab} \in E, \forall i=1, \dots, C, \forall k=1, \dots, K.$$

$$l_{ab}^i = l_{ab}^i[x], \forall a, b \in N, a, b \notin \{p, q\}, e_{ab} \in E, \forall i=1, \dots, C.$$
- 6: 设置 $l_{ab}^i[x+1] = l_{ab}^i \forall a, b \in N, e_{ab} \in E, \forall i=1, \dots, C$.
- 7: end
- 8: 使用文献[2]中的方法给每个流确定一条路径
- 9: 根据如下限制条件解优化问题

由第8行获得 $l_{ab,i}^k, l_{ab}^i = l_{ab}^i[X], \forall a, b \in N, e_{ab} \in E, \forall i=1, \dots, C, \forall k=1, \dots, K$.

3 性能分析

本节通过仿真来分析所提出的跨层公平资源分配机制的性能。网络域的大小为 $550* \times 550*$ 。每次仿真均是 20 次不同随机场景结果的平均。在每个场景，MRMC-WMN 包括 10 个随机均匀分布的无线 Mesh 路由器，其中两个分布在仿真区域边缘的节点作为网关。节点的通信范围和干扰范围分别为 250m 和 450m。每个路由器配置有 3 个射频；网络中有 6 个可用正交信道。设置参数： $\alpha=2$ ， $c_p=54$ 。在每一个场景，网络中有 5 个内部流和 5 个外部流，每个流的速率为 2Mbps。对每个内部流，源节点和目的节点都是从非网关节点中随机选取。每个外部流是从随机选择的非网关节点传输到网关节点。

3.1 逻辑拓扑设计、射频和信道分配

图 1 是本文所提出机制的逻辑拓扑设计、射频和信道分配结果。图中虚线表示没有建立逻辑链路的边；实线表示逻辑链路，边上的数字表示分配给该链路的信道。其中节点 a 和节点 b 之间有 2 条逻辑链路。路由器上的黑点表示射频，其中有些逻辑链路需要共享一个射频，如逻辑链路 (d,h) 和 (d,e) 、 (b,d) 和 (b,a) 等。

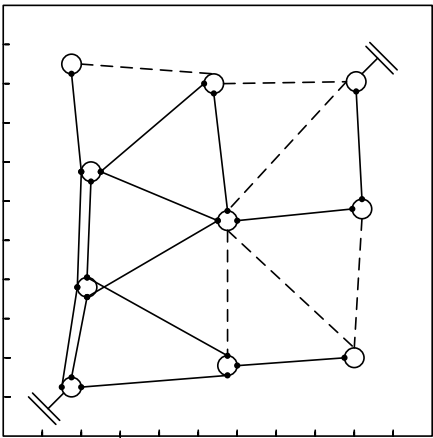


图 1 逻辑拓扑设计、射频和信道分配结果

3.2 近似最优解的性能

在这一部分，将所提出的 TSILS 算法性能与分支定界法和 ILS^[3]算法作对比。令 ω_{\min}^* 表示由分支定界法获得的最优解， $\omega_{\min}[x]$ 表示 ILS 算法和 TSILS 算法在第 x^{th} 次迭代所获得的解。使用 $(\omega_{\min}^* - \omega_{\min}[x]) / \omega_{\min}^*$ 作为优化偏差^[2]。图 2 表示当节点和流数目相等，且数目为 10 和 20 时，平均优化偏差和迭代次数的关系。由图可知，平均优化偏差随着迭代次数的增加而减少。当网络规模增大（节点数目增加）时，获得近似最优解的迭代次数增加。因为，网络规模增大使得边增加，所以需要迭代更多次数来给边重新分配信道。由图可知，TSILS 算法比 ILS 算法更快获得近似最优解，因为由算法 1 获得的信道分配初始解具有最小网络干扰，且初始

解的带宽远远大于 ILS 中的单信道逻辑拓扑带宽。当节点和流数目为 10 和 20 时, TSILS 能够分别在 60 和 130 次迭代后获得不大于 3% 的优化偏差。

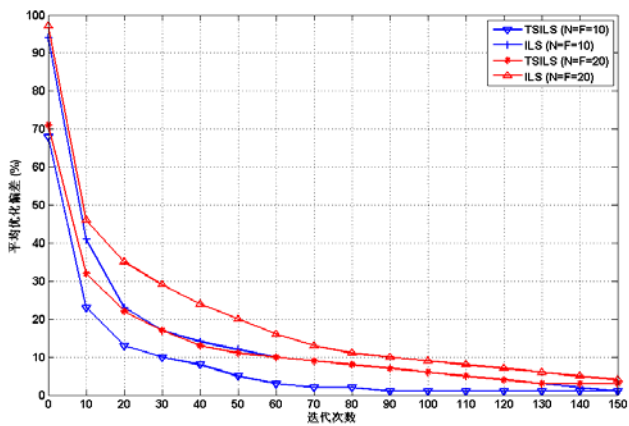


图 2 平均优化偏差和迭代次数的关系

3.3 网络资源对性能的影响

本节估计射频和信道数目对系统性能的影响, 所仿真的 MRMC-WMNs 包括 10 个 Mesh 路由器和 10 个流的, TSILS 算法的迭代次数为 60。考虑两个性能参数: 网络吞吐量和公平性。网络吞吐量表示所有逻辑链路累计的链路层带宽, 即 $\sum_{e_{ab} \in E, i=1, \dots, C} c_{ab}^i$ 。公平性采用文献[17]中的定义: $(\sum_{e_{ab} \in E, i=1, \dots, C} (\varepsilon c_{ab}^i - \beta_{ab}^i)^2) / (|L| \sum_{e_{ab} \in E, i=1, \dots, C} (\varepsilon c_{ab}^i - \beta_{ab}^i)^2)$ 。图 3 和图 4 分别表示网络资源对吞吐量和公平性的影响, 其中 I 表示射频数目。由图可知: (1) 网络吞吐量和公平性随着信道数目的增加而平稳的增加; (2) 相对于单射频和单信道而言, 本文所提出的设计能够获得更高的吞吐量和更好的公平性; (3) 在信道数目一定的情况下, 当射频数目从 3 增加到 4 时, 吞吐量的增加和公平性的改善很少。所以, 给每个节点配置 2 到 3 个射频最有效和经济。

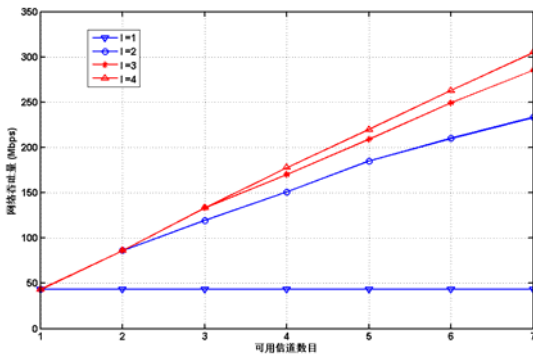


图 3 网络资源对吞吐量的影响

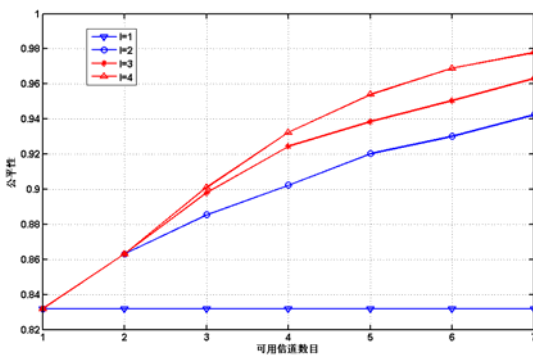


图 4 网络资源对公平性的影响

4 结论

本文提出一种 MRMC-WMNs 中的跨层公平资源分配机制。首先对逻辑拓扑设计, 信道

分配、路由和带宽分配进行了数学建模，然后将该跨层公平资源分配问题描述成一个整数线性规划(ILP)的优化问题，并提出一种近似优化的后启发式 TSILS 算法来求解该问题。所提出的机制能够在链路间最大-最小公平性限制下最大化网络带宽。仿真结果表明，所提出的资源分配机制能够保证链路间的公平性，且相比 ILS 算法具有更好的收敛性和更小的优化偏差。同时研究了网络资源对吞吐量和公平性的影响，能够给网络资源配置提供参考。今后，我们将继续研究在 MRMC-WMNs 中为不同类型的业务提供服务质量保证。

参 考 文 献

- [1] P. Bahl, A. Adya, J. Padhye, and A. Wolman, "Reconsidering wireless systems with multiple radios," *ACM Computer Commun. Rev.*, vol. 34, pp. 39–46, Oct. 2004.
- [2] A. Hamed Mohsenian-Rad and Vincent W. S. Wong, "Joint Logical Topology Design, Interface Assignment, Channel Allocation, and Routing for Multi-Channel Wireless Mesh Networks," *IEEE Trans. on Wireless Commun.*, vol.6, no.12, pp. 4432-4440, Dec. 2007.
- [3] Chen L, Low S H, Chiang M, et al Cross-layer congestion control, routing and scheduling design in Ad Hoc wireless networks [C] // *Proceedings of the 25th Conference on Computer Communications Barcelona, Spain: IEEE*, 2006: 1-13.
- [4] Ian F. Akyildiz and X. D. Wang, "Cross-Layer Design in Wireless Mesh Networks," *IEEE Trans. on Vehicular Technology*, vol. 57, no. 2, pp. 1061-1076, March 2008.
- [5] M. Chiang, S. H. Low, A. R. Calderbank and J. C. Doyle, "Layering as Optimization Decomposition: A Mathematical Theory of Network Architectures," in *Proc. IEEE*, Vol. 95, Issue 1, pp.255-312, Jan. 2007.
- [6] A. H. Mohsenian Rad and V. W. S.Wong, "Joint optimal channel assignment and congestion control in multi-radio wireless mesh networks," in *Proc. IEEE ICC'06*, June 2006.
- [7] A. H. Mohsenian Rad and V. W. S. Wong, "Joint optimal channel allocation, interface assignment, and MAC design for multi-channel wireless mesh networks," in *Proc. IEEE Infocom'07*, Anchorage, AK, May 2007.
- [8] M. Alicherry, R. Bhatia, and L. Li, "Joint channel assignment and routing for throughput optimization in multi-radio wireless mesh networks," in *Proc. ACM MobiCom'05*, Sept. 2005.
- [9] X. Meng, K. Tan, and Q. Zhang, "Joint routing and channel assignment in multi-radio wireless mesh networks," in *Proc. IEEE ICC*, June 2006.
- [10] M. Kodialam and T. Nandagopal, "Characterizing the capacity region in multi-radio multi-channel wireless mesh networks," in *Proc. ACM MobiCom'05*, Sept. 2005.
- [11] S. Iyer, S. Bhattacharyya, N. Taft, and C. Diot, "An approach to alleviate link overload as observed on an IP backbone," in *Proc. IEEE Infocom'03*, April 2003.
- [12] K. Papagiannaki, S. Moon, C. Fraleigh, P. Thiron, F. Tobagi, and C. Diot, "Analysis of measured single-hop delay from an operational backbone network," in *Proc. IEEE Infocom'02*, June 2002.
- [13] D. P. Bertsekas and R. Gallager, *Data Communications*, 2nd ed. Prentice Hall, 1992.
- [14] J. Mo and J. Walrand, "Fair end-to-end window-based congestion control," *IEEE/ACM Trans. Networking*, vol. 8, pp. 556–567, Oct. 2000.
- [15] D. P. Bertsekas, *Nonlinear Programming*, 2nd ed. Athena Science, 2004.
- [16] Capone and M. Trubian, "Channel assignment problem in cellular systems: a new model and a tabu search algorithm," *IEEE Trans. on Veh. Tech.*, vol. 48, no. 4, pp. 1252-1260, July 1999.

[17] R. Jain, W. Hawe, and D. Chiu, “A quantitative measure of fairness and discrimination for resource allocation in shared computer systems,” Tech. Rep. DEC-TR-301, Sept. 1984.

作者简介:

黄鑫（1984 年-），女，华南理工大学电子与信息学院在读博士研究生，主要研究领域为无线 Mesh 网络的跨层优化、资源分配等。

冯穗力（1955 年-），男，华南理工大学电子与信息学院教授、博士生导师。主要研究领域为：信息网络理论与技术、多媒体信号传输与处理及数字信号处理在通信中的应用等。

庄宏成，男，华为技术有限公司资深工程师。

DF协作中继网络基于最小能量成本的中继选择和功率控制策略

柯 峰 冯穗力

(华南理工大学电子与信息学院, 广东 广州 510640)

摘 要: 本文针对解码-转发 (DF) 协作中继网络, 提出了一种基于最小能量成本的中继选择和功率分配策略, 一方面考虑减少每次协作传输的能量消耗, 另一方面通过给节点剩余能量进行定价平衡网络中每个节点的能量消耗。通过图解法可以得到源节点和中继节点的成本最小的最优功率分配解, 降低了算法的复杂度。仿真结果表明, 本文算法与其他算法相比, 可以有效延长整个协作中继网络的寿命。

关键词: 协作中继; 解码-转发; 中继选择; 功率分配

Relay Selection and Power Allocation for Decode-and-Forward (DF) Cooperative Network Based on Minimum Energy Cost

Ke Feng Feng Sui li

(School of Electronic and Information Engineering, South China University of Technology, Guangzhou 510640, China)

Abstract: In this article, a relay selection combined with power allocation scheme for decode-and-forward cooperative network is proposed based on energy pricing which takes into account both reducing the energy consumption of each transmission and balancing the usage of every node's energy in the network. The optimal power allocation solution based on minimum energy cost between the source node and the relay node is obtained by graphic method which reduces the computational complexity. Simulation results show that the lifetime of the cooperative network can be significantly prolonged using the proposed scheme compared with other schemes.

Keywords: cooperative relaying; decode-and-forward; relay selection; power allocation

1 引言

无线网络通过协作中继传输技术如放大-转发(AF), 解码-转发(DF)等协议的实现, 可以通过空间分集实现通信链路质量的有效改善^[1], 中继选择和功率分配是影响协作中继网络性能的重要因素。在文献[2]中, 证明了每次只选择一个协作中继可以有效减少网络信令开销, 同时

基金项目: 粤港科技关键领域突破项目 (20060104-2); 广东省自然科学基金博士启动基金 (06300097)

与每次选择多个协作中继相比,性能并没有下降。基于误符号率的分析,一个基于解码-转发协作协议的功率分配策略在文献[3]中被提出。J.Luo 等人在即时功率约束下试图得到整体中断概率最小的功率分配解^[4]。

但为了延长网络的使用寿命,我们不仅要降低每次传输的能量消耗,还要考虑平衡网络中每个节点的能量消耗。在本文中,我们将网络中的节点建模为能量卖家,可根据剩余能量的多少对能量进行定价。将用户建模为买家,根据能量成本最小准则(MIC)选择最优的协作伙伴和在源节点和中继节点之间进行功率分配。根据以上这些设定,我们提出了针对解码-转发(DF)协作中继网络的一种基于能量定价的中继选择和功率分配策略,并通过仿真与其他策略进行了比较。

2 系统模型和问题描述

A. 系统模型

假定协作中继网络由一个源节点(S),一个目的节点(D),以及 K 个分布在不同位置的中继站构成。中继集合 Ω 以表示,且 Ω 中的每个中继可以协作源节点(S)的传输。假定在每个数据块传输期间,信道条件保持稳定。以 x_S 表示源节点(S)待发送的数据符号, $g_{S,D}, g_{S,R(k)}, g_{R(k),D}$ 分别表示 $S \rightarrow D$ 链路, $S \rightarrow R(k)$ 链路以及 $R(k) \rightarrow D$ 链路的信道衰落系数。此外,以 W 表示中继网络的传输带宽。不失一般性,假定 $S \rightarrow D$ 链路, $S \rightarrow R(k)$ 链路以及 $R(k) \rightarrow D$ 链路上叠加的高斯白噪声 $n_{S,D}, n_{S,R(k)}, n_{R(k),D} \sim CN(0, \sigma^2)$ 。此外,分别以 $\xi_{S,D} = g_{S,D}^2 / \sigma^2$, $\xi_{S,R(k)} = g_{S,R(k)}^2 / \sigma^2$ 和 $\xi_{R(k),D} = g_{R(k),D}^2 / \sigma^2$ 表示 $S \rightarrow D$ 链路, $S \rightarrow R(k)$ 链路以及 $R(k) \rightarrow D$ 链路增加单位发射功率所增加的信噪比。

如果 $S \rightarrow D$ 链路的信道条件足够好,源节点(S)可以直接通过 $S \rightarrow D$ 链路发送符号到目的节点(D)。否则,源节点 S 会从中继集合 Ω 中选择一个最优中继 R^{opt} 进行协作传输。因而,有两个问题需要得到解决。第一,如果从中继集合 Ω 中选择最优的中继 R^{opt} 。第二,发射功率为多少。

假定源节点的数据块的所需传输速率为 I_b (比特/s),分配给数据块的传输时隙长度为 T_b 。在不需要中继节点的帮助下,源节点(S)实现数据块传输所需的发射功率 p_S^{dir} 满足下式:

$$I_b = W \log_2(1 + \xi_{S,D} p_S^{dir}) \quad (1)$$

从式(1)可以得到:

$$p_S^{dir} = (2^{I_b/W} - 1) / \xi_{S,D} \quad (2)$$

如果 $S \rightarrow D$ 链路的信道质量不好或者源节点(S)的最大发射功率不够大,源节点需要从中继集合 Ω 中选择一个中继 $R(k)$ 实现协作传输并决定源节点(S)和中继节点 $R(k)$ 的发射功率 $p_S, p_{R(k)}$ 。协作传输机制基于时分复用,即每个数据块的传输被分为两个时隙,每个时隙的长度为 $T_b/2$ 。在第一个时隙内,源节点(S)向目的节点(D)和 Ω 中的中继节点广播符号 x_S 。于是可以得到目的节点(D)和中继 $R(k)$ 接收到的符号如下所示:

$$y_{S,D} = \sqrt{p_S} g_{S,D} x_S + n_{S,D} \quad (3)$$

$$y_{S,R(k)} = \sqrt{p_S} g_{S,R(k)} x_S + n_{S,R(k)} \quad (4)$$

然后, 中继 $R(k)$ 对接收到的信号进行解调, 解码, 对估计得到后的符号 $\hat{x}_{R(k)}$ 重新编码, 调制, 在第二个时隙以 $p_{R(k)}$ 发射功率进行发射。

在目的节点(D), 通过对来自源节点和中继 R^{opt} 的信号进行最大比合并(MRC)后得到最后的数据。整个链路的可达速率由 $S \rightarrow R_k$ 链路和 $R_k \rightarrow D$ 链路的信噪比(SNR)的最小值决定。即如下式所示:

$$I_b \leq (1/2)W \log_2(1 + \min\{\xi_{S,R(k)}p_S, \xi_{S,D}p_S + \xi_{R(k),D}p_{R(k)}\}) \quad (5)$$

不失一般性, 我们假定源节点和 Ω 中所有中继节点的最大发射功率满足 $p_S^{\max} = p_{R(1)}^{\max} = \dots = p_{R(K)}^{\max} = p_{\max}$ 。

从(5)式, 我们可以得到建立有效 $S \rightarrow R(k) \rightarrow D$ 链路的源节点(S)的发射功率的下界 p_S^{low} , 如下式所示:

$$p_S^{low} = (2^{2I_b/W} - 1) / \xi_{S,R(k)} \quad (6)$$

此外, 我们还可以得到要实现有效协作传输时源节点的发射功率的上限 $p_S^{up} = p_{\max}$ 。此外, 我们可以得到中继 $R(k)$ 的发射功率必须在 $[0, p_{\max}]$ 的范围之内。

B. 问题的描述

可以将用户看成一个买家, 通过向源节点和中继节点购买能量, 完成数据块的传输,

并在保持传输速率 I_b 的同时尽可能降低消耗的成本。源节点(S)和 Ω 中所有中继节点都可以看成卖家, 根据自身剩余的能量多少对能量进行定价。在完成数据块传输的时隙内, 源节点和中继节点所消耗的能量分别为 $e_S = p_S T_b / 2$ 和 $e_{R(k)} = p_{R(k)} T_b / 2$, 于是可以得到成本函数如下所示:

$$C_{S,R(k)}(p_S, p_{R(k)}) = (1/2)T_b(\rho_S p_S + \rho_{R(k)} p_{R(k)}) \quad (8)$$

式(8)中, $\rho_S, \rho_{R(k)}$ 分别为源节点(S)和中继节点 $R(k)$ 的单位能量的价格。

以 $(p_S^*, p_{R(k)}^*)$ 表示当中继 $R(k)$ 被选择时取得最小成本的最优功率分配解, 于是可以得到最优的中继如下所示:

$$R^{opt} = \arg \min_{R(k)} C_{S,R(k)}(p_S^*, p_{R(k)}^*) \quad (9)$$

如果 $C_{S,R^{opt}}(p_S^*, p_{R^{opt}}^*) > C_S^{dir}$, 则对用户来说, 没有必要中继进行协作传输, 其中 $C_S^{dir} = T_b \rho_S p_S^{dir}$ 代表用户直接传输所需耗费的成本。

由上可以得到求解中继 $R(k)$ 的最优功率分配 $(p_S^*, p_{R(k)}^*)$ 问题可以数学建模为一个典型的线性规划问题, (10) 式定义了问题的优化目标和需满足的约束条件:

$$\begin{aligned} (p_S^*, p_{R_k}^*) = \arg \min_{p_S, p_{R_k}} & (\rho_S p_S + \rho_{R_k} p_{R_k}) \\ \text{s.t.} & \begin{cases} \xi_{S,D} p_S + \xi_{R(k),D} p_{R(k)} \geq 2^{2I_b/W} - 1 \\ p_S^{low} \leq p_S \leq p_{\max} \\ 0 \leq p_{R(k)} \leq p_{\max} \end{cases} \end{aligned} \quad (10)$$

C. 能量的定价策略

为了延长网络的使用寿命, 剩余能量少的节点必须对能量定一个高一点的价格, 反之剩余能量多的节点应对能量定一个低一点的价格。在本文中, 我们假定节点定的价格与剩余能

量倒数的一次阶或者二次阶成正比。

即如下式所示：

$$\rho_{R(k)} = \beta \left(\frac{E_{R(k)}^{tot}}{E_{R(k)}^{rem}} \right)^m \quad m=1,2 \quad (11)$$

式中， $E_{R(k)}^{tot}$ 表示中继 $R(k)$ 的初始能量和， $E_{R(k)}^{rem}$ 表示数据传输前中继 $R(k)$ 的剩余能量， m 为 1 或 2。

3 最优功率分配问题的图解法

为了得到最优的功率分配解 $(p_S^*, p_{R(k)}^*)$ ，定义以下三个相关函数：

$$f(x, y) = (W/2) \log_2(1 + \xi_{S,D}x + \xi_{R(k),D}y) \quad (12)$$

$$h_1(x) = (2^{2I_b/W} - 1 - \xi_{R(k),D}x) / \xi_{S,D} \quad (13)$$

$$h_2(x) = (2^{2I_b/W} - 1 - \xi_{S,D}x) / \xi_{R(k),D} \quad (14)$$

如图 1 所示，直线 $S_i (i=1,2,3)$ 为不同 I_b 取值下的信噪比 (SNR) 等值线 $\xi_{S,D}p_S + \xi_{R(k),D}p_{R(k)} = 2^{2I_b/W} - 1$ ，图中阴影区域为矩形 ABCD 与 S_i 的右上半平面的交集，由此可见阴影区域为式 (11) 定义的优化问题的可行域。由图可见， I_b 越大，直线 S_i 位置越靠上，阴影区域即可行域的面积越小。定义两个比值参数 $\eta_1 = \rho_S / \rho_{R(k)}$ ， $\eta_2 = \xi_{S,D} / \xi_{R(k),D}$ ，直线 C_1 和 C_2 分别为当 $\eta_1 \geq \eta_2$ 和 $\eta_1 < \eta_2$ 时的成本等值线 $\rho_S p_S + \rho_{R(k)} p_{R(k)} = 2C_{S,R(k)} / T_b$ ，由图 1 可见，直线 C_1 和 C_2 越靠下，成本 $C_{S,R(k)}$ 越低。则容易知道，从下往上移动直线 C_1 和 C_2 ，与可行域阴影区域的第一个交点即为最优的功率分配解 $(p_S^*, p_{R(k)}^*)$ 。

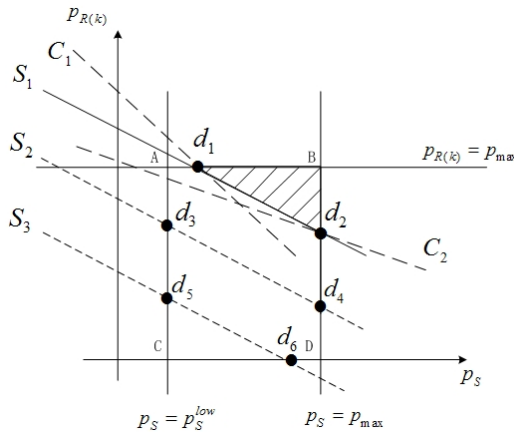


图 1 最优功率分配 $(p_S^*, p_{R(k)}^*)$ 的图解法

当 $I_b > f(p_{\max}, p_{\max})$ 或者 $p_S^{low} \geq p_{\max}$ ，由图 1 可见，当这种情况发生时，可行域的面积为零，意味着不存在满足式 (11) 的有效解，因此这种情况下， $R(k)$ 将不可能为用户所选择。

当 $I_b < f(p_S^{low}, 0)$ 时，由图 1 可见，当这种情况发生时，可行域为整个矩形 ABCD，最佳

功率分配点为 C，由于 $p_{R(k)}^* = 0$ ，意味着中继 $R(k)$ 将不可能为用户所选择。

当 $f(p_S^{low}, 0) \leq I_b \leq f(p_{max}, p_{max})$ ，且 $p_S^{low} < p_{max}$ ，由图 1 可知，此时，阴影部分的面积不为零，于是可以得到最佳的功率分配解，如下所示：

A. 当 $\eta_1 > \eta_2$ 时

如图 1 所示，此时成本等值线 C_1 的斜率大于信噪比(SNR)等值线 S_i 的斜率，当 $f(p_S^{low}, p_{max}) \leq I_b \leq f(p_{max}, p_{max})$ 时，最佳功率分配点如 d_1 所示，位于线段 AB 之上。当 $f(p_S^{low}, 0) \leq I_b \leq f(p_S^{low}, p_{max})$ 时，最佳功率分配点如 d_3 或 d_5 所示，位于线段 AC 之上。于是可以得到最优功率分配向量如下所示：

$$(p_S^*, p_{R(k)}^*) = \begin{cases} (h_1(p_{max}), p_{max}) & f(p_S^{low}, p_{max}) < I_b \leq f(p_{max}, p_{max}) \\ (p_S^{low}, h_2(p_S^{low})) & f(p_S^{low}, 0) \leq I_b \leq f(p_S^{low}, p_{max}) \end{cases} \quad (15)$$

B. 当 $\eta_2 > \eta_3$ 时

此时成本等值线 C_2 的斜率小于信噪比(SNR)等值线 S_i 的斜率，当 $f(p_{max}, 0) \leq I_b \leq f(p_{max}, p_{max})$ 时，最佳功率分配点如 d_2 或 d_4 所示，位于线段 BD 之上。当 $f(p_S^{low}, 0) \leq I_b \leq f(p_{max}, 0)$ 时，最佳功率分配点如 d_6 所示，位于线段 CD 之上。于是可以得到最优功率分配向量如下所示：

$$(p_S^*, p_{R(k)}^*) = \begin{cases} (p_{max}, h_2(p_{max})) & f(p_{max}, 0) < I_b \leq f(p_{max}, p_{max}) \\ (h_1(0), 0) & f(p_S^{low}, 0) \leq I_b \leq f(p_{max}, 0) \end{cases} \quad (16)$$

4 仿真与分析

在本节中，对本文基于最小成本(MIC)准则的中继选择和功率分配策略基于 MATLAB6.5 平台进行性能仿真，与文献[6]中的最小能量消耗(MIE)和最大剩余能量(MARE)策略进行比较。考虑如下场景，在一个矩形区域内均匀分布 25 个节点，相邻节点间距为 10 米，路径损失因子为 2.5，节点之间的信道为叠加高斯白噪声的瑞利衰落信道，不失一般性，假定所有信道的噪声功率 $\sigma^2 = 10^{-8}$ ，并假定每个时隙内，都只有一个随机选择的源节点发送信息，目的节点因为随机选择。时隙长度 $T_b = 10^{-3}$ s。每个节点的初始能量为 1 焦耳，最大发射功率为 0.1W。带宽为 1M 赫兹，图 2 为各种策略下，不同数据传输速率下协作中继网络寿命的对比。由图可见，本文算法与其他算法相比，可以明显改善网络的使用寿命，尤其在速率低的情况下更加明显。

5 结论

本文提出了一种 DF 协作中继网络下基于最小能量成本(MIC)准则的中继选择和功率分配策略，一方面考虑减少每次协作传输的能量消耗，另一方面通过给节点剩余能量进行定价平衡网络中每个节点的能量消耗。通过图解法可以得到源节点和中继节点的成本最小的最优功率分配解，降低了算法的复杂度。仿真结果表明，本文算法与其他算法相比，可以有效延长整个协作中继网络的寿命，尤其在速率低的情况下更加明显，因而具有较好的实用价值。

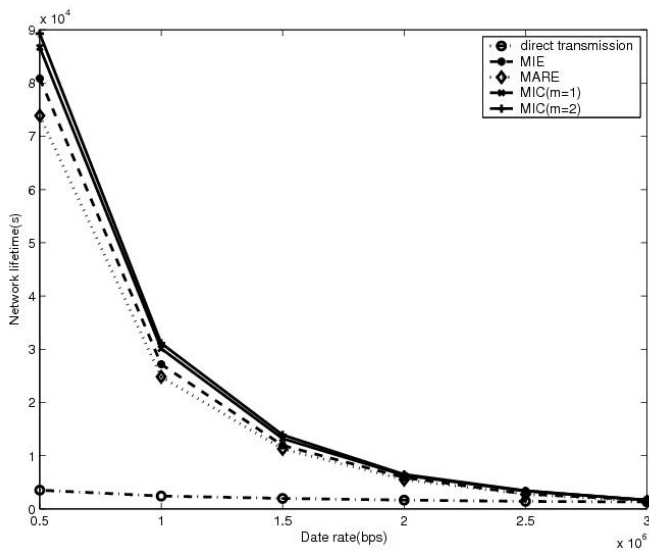


图2 不同数据传输速率下网络的寿命对比

参 考 文 献

- [1] J. N. Laneman, D. N. C. Tse and G. W. Wornell.Cooperative diversity in wireless networks: efficient protocols and outage behavior[J]. IEEE Trans.Inform. Theory, vol.22, pp.3062 - 3080, Dec. 2004.
- [2] A.Bletsas, A.Khisti, D.Reed, and A. Lippman.A simple cooperative diversity method based on network path selection[J]. IEEE Trans. Wireless Commun, vol.24, pp.659 - 672, March. 2006.
- [3] W.Su, A. K. Sadek, and K.J.R.Liu,.SER performance analysis and optimum power allocation for decode-and-forward cooperation protocol in wireless networks[C]. Proc. IEEE Wireless Communications and Networking Conference.vol.2, pp.984 - 989, March. 2005.
- [4] J.Luo et al.Decode-and-forward cooperative diversity with power allocation in wireless networks[J]. IEEE Trans.Wireless Commun.,vol.6, pp.793 - 799, Mar. 2007.
- [5] Wan-Jen.Huang, Hong, and Y.-W.P. et all.Lifetime maximization for amplify-and-forward cooperative networks[J]. IEEE Trans. Wireless Commun.,vol.7, pp.1800 - 1805, May. 2008.
- [6] Zhong.Zhou,Shengli.Zhou and Jun-Hong.Cui.Energy-efficient cooperative communication based on power control and selective single-relay in wireless sensor networks[J]. IEEE Trans. Wireless Commun., vol.7, pp.3066- 3078, August. 2008.
- [7] Y. Chen and Q. Zhao,.On the lifetime of wireless sensor networks[J]. IEEE Commun. Lett., vol.9, pp.976 - 978, Nov. 2005.
- [8] Landsburg and Steven E., Price theory and applications[M], Australia:South-Western Pub.,2002.

作者简介:

柯峰(1975-), 男, 汉族, 江苏无锡人, 博士, 华南理工大学电子与信息学院, 讲师, 主要从事宽带移动通信, 数字集成电路方面的研究。

冯穗力(1955-), 男, 汉族, 广东省梅州人, 博士, 华南理工大学电子与信息学院, 教授, 博导, 主要从事宽带移动通信, 下一代互联网络方面的研究。

在WSNs中基于覆盖与连通的分簇路由协议的改进

朱丽娟 王敬国 林晓辉 王 晖

(深圳大学 信息工程学院, 广东 深圳, 518060)

摘 要: 在无线传感器网络中, 为提供区域监控的有效性, 网络的连通及覆盖至为关键。基于这个要求, 对传统的分簇路由协议 LEACH 的簇头选举阶段进行了改进, 提出了 LEACH-CC 协议, 并通过 NS2 对其进行仿真和比较。仿真结果表明, 有效的提高了网络的连通和覆盖, 增长了网络的生命周期。

关键字: 无线传感器网络; 覆盖与连通; LEACH; LEACH-CC

On the Cluster-Based Routing Protocol of Wireless Sensor Networks: Coverage, Connectivity and the Improvement

Zhu Li Juan Wang Jing guo Lin Xiao hui Wang Hui

(College of Information Engineering, Shenzhen University, Shenzhen 518060, China)

Abstract: Connectivity and coverage of wireless sensor network (WSN) are two critical issues in the provision of effective environment surveillance. To meet these requirements for a traditional cluster-based WSN, in this paper, we make some improvements on cluster head selection mechanism by proposing a routing protocol called LEACH-CC. We also perform extensive simulations on LEACH-CC. The results demonstrate that the proposed protocol can effectively improve the network connectivity and coverage. Additionally, the network lifetime has also been prolonged.

Keywords: WSNs; Coverage and Connectivity; LEACH; LEACH-CC

无线传感器网络^[1]是由大量微型传感器节点组成的, 具有自组织能力自动的进行配置和管理, 快速的按照预设的协议组织成一定目的的网络进行信息的采集和发送, 并且随着环境的多变而动态的调整拓扑结构, 具有一定的鲁棒性和容错性, 而且能与传统的网络进行有效的互联, 具有广泛的应用前景。例如军事、环境监测和预报、健康管理、智能家居和城市交通等等。

由于无线传感器网络是针对监测的物理环境进行数据的有效采集, 所以对整个环境的覆盖是一个重要的前提。只有满足的全覆盖或者高覆盖, 才能对其中的温度、湿度等一系列感兴趣的信息进行捕捉采集, 才能更详细更全面的掌握客观世界的讯息。而无线传感器网络大多数都是随意抛洒节点在监测区域中, 所以很难有规划的进行全覆盖的预置。传感器节点在通常通过布置之后的定位消息, 能量, 周围邻居节点信息等因素来考虑覆盖问题, 当然大量

的研究工作人员对此做了各方面的研究。

本论文基于覆盖与连通的思想,对无线传感器网络中第一个分簇路由协议 LEACH 协议进行改进。以下就是进行的主要工作。

1 相关工作

LEACH^[2]协议是无线传感器网络中所提出的第一个分层路由协议。它的工作过程以“轮”来计算。每一轮分为簇的建立阶段和稳定的数据传输阶段。在簇的建立阶段,每个节点产生一个随机数,随机产生簇头,动态的形成簇。在数据传输阶段,各节点按照簇头分配的 TDMA 消息在各自的时间槽向簇头发送采集来的消息,簇头节点在收集各节点发来的消息后对其进行数据融合将最终的消息发给基站(sink)节点。在 LEACH 协议中,节点不需要知道全局信息,也不需要复杂的控制信息,增添了网络能量均衡控制,保障了节点的能耗均匀,增长了网络的生命周期。但由于其随机的动态的分簇过程,没有充分的考虑到簇头节点的能量消耗过快,以及簇头在整个网络中的分布问题,常导致整个网络分布不均,个别节点能量快速消耗导致网络空洞等问题。

针对 LEACH 协议的优点和不足,大量的研究人员对其进行研究和改进。Younis^[3]等人将节点的剩余能量考虑到初始概率的竞争过程中,并不是产生一个随机数,提出了混合能量有效分布的 HEED 算法,优化了全网的能量消耗,使得网络的分布更合理。Lindsey 等人提出了能量有效汇聚 PEGASIS^[4]路由协议,使簇头节点形成一条链路传输数据,每一轮只有离基站最近的一个簇头直接与基站通信,减少直接与基站通信的节点数目,减少长距离的传输,有效的节约能量,适用于大规模远距离的网络。Ya^[5]等人提出的 GAF 算法,则是根据节点的地理位置先分簇,把监测区域划分成虚拟单元格,将节点按照位置信息划入相应的单元格,在每个单元格中定期选举产生一个簇头节点,这就需要传感器节点高的定位精度。

对于网络的连通与覆盖问题,研究人员通过很多图论和数学的证明过程,并且结合无线传感器网络特殊的特点进行研究和分析。在文献[6]中,提出了一个基于可信度的密度控制算法,在自由的网络空间中,满足较长的生命周期、鲁棒的传感覆盖。在这个协议中只需要一组节点的子集就可以达到所需传感覆盖要求的网络模式,其余节点进入休眠状态;文献[7]从实际的节点功率的角度出发,节点的发射功率不能太小,影响网络的联通度,节点功率也不能太大,否则会产生严重的信号干扰;Sachin and Mani^[8]等人在需要监测区间,基于一个暴露模型去决定全覆盖的情况下所需的传感器节点数;Jorge^[9]等人针对移动 WSNs 提出了自适应、分布式以及异步的覆盖算法;Xiaojiang 和 Fargo^[10]等人将网络分为不同的区域,针对不同的网络有不同的覆盖算法,也就是满足不同的覆盖率。

2 网络模型

所有传感器节点随机的分布在监控区域 S 中。监控区域相对于传感器节点的传感半径足够大;分布密度足够的大,使得在所有节点处于工作状态时,能够覆盖整个区域,节点至少能与一个邻居节点直接连通。整个节点模型有以下假设:

- (1) 所有节点都是同构的,平等的,都能直接与基站通信;

- (2) 所有节点在布置在二维的平面内，布置后，位置固定，不能移动；
- (3) 节点采用布尔感知模型，每个节点的传感半径为 r ，节点的传感覆盖范围是以 r 为半径的圆，传感覆盖面积是 πr^2 。在传感半径以内的信息可以被节点感知，以外则不能；
- (4) 通过调节节点的发射功率而获得所需要的不同的通信半径；
- (5) 节点不需要精确的定位信息。

3 问题分析

LEACH 的一个优点就是不需要太多的上层控制信息，在组网的过程过方便快捷，也正是能量有限计算能力有限的传感器所需要的。那么采用传统的覆盖和连通的理论，就需要通过 GPS 对其进行精确的定位，相互交换邻居节点乃至整个网络的信息，那么就会对传感器造成很大的负担。所以我们考虑节点感知半径以内的邻居节点数目来衡量覆盖与连通的问题。

文献[11]主要根据节点的邻居节点情况来判断节点的冗余性。首先从节点扇形区和扇形角度，引入图论的多个定理极其推论得出两个重要结论。第一，在无位置信息和方向信息的精确定位的情况下，设节点的邻居节点数为 n ，事件 A 是节点的覆盖区域被这 n 个邻居节点完全覆盖，则 A 发生的概率 $Pr\{A\}$ 是：

$$1 - 0.609^{n-1} \leq P_r\{A\} \leq 1 - n0.0609^{n-1} + \varepsilon \tag{1}$$

其中 $\varepsilon = n(n-1)(0.276)^{n-1} / 2$ 。

第二，节点的传感范围被这 n 个邻居节点覆盖的百分比不小于：

$$1 - 0.609^n - \frac{1}{2}n0.06^n \left[(-1)^n \frac{4}{3} - \frac{2}{n} \right] \tag{2}$$

计算结果得出，节点如果成为完全冗余节点，监测区域覆盖率为 100% 的代价非常大。当节点的邻居节点数达到 11 时，它成为完全冗余节点的概率只有 92.28%。同时也可以看出，当节点的邻居节点数达到 5 时，90% 以上的监测区域会被邻居节点覆盖。

文献[12]是关于网络连通情况下节点的最佳邻居节点数的研究。Xue 得出在节点数为 n 的网络中，随着 n 的增加，最佳邻居节点数不是一个恒定值，事实上这个数是与 $\Theta(\log n)$ 方式递增。这个与先前在七、八十年代提出的“魔法数字”认为 6 和 8 是最优邻居节点数相适应。但是这个结果不适应于大型网络。文章还通过一系列严谨的理论和网络仿真得到，如果每个节点的邻居节点数少于 $0.074 \log n$ ，则整个网络将处于不连通状态；如果每个节点邻居节点数多于 $5.1774 \log n$ 时，网络连通的概率趋于 1。由此可得，对于小规模的网络来说，8 或 6 个邻居节点确实可以以很高的概率保证网络连通。

文献[13]同样从节点的扇形角度出发，引入数学公式和图论定理证明了网络中连通和覆盖的关系。假设当所有节点同构的，传感半径是 r ，通信半径是 R ，当 $R \geq 2r$ 时，网络全覆盖的条件下，就可以满足网络的全连通。同理，在 $k(k > 1)$ 重覆盖 k 重连通的情况下， $R \geq 2r$ 也是在 k 重覆盖的情况下，保证 k 重连通的充分条件。而且随着最初布置节点和节点的传感半径的增加，将使更多的节点成为休眠节点。当然文中也分析了节点异构的情况下，覆盖与连通的问题。

综上所述，把邻居数大于等于 5 的节点定义为冗余节点，这些节点有机会进入休眠状态，小于 5 的节点将一直处于活动状态，就可以保证网络的覆盖率在 90% 以上。而此时通信半径 $R \geq 2r$ 的情况下就可以保证在高覆盖的条件下的全连通了。

4 改进的LEACH-CC协议

LEACH-CC 协议中，簇头即要收集簇内节点采集来的信息，同时要对其进行数据融合，将数据包发送给远距离的基站。所以，簇头节点往往要比普通节点消耗能量快，也会因为能量的快速消耗而过早的死亡，造成网络的空洞，影响对整个监测区域的覆盖和连通。所以，我们采用让邻居节点数大于或等于 5 的冗余节点才有机会参与簇头的选举。这样选举出来的簇头都是被其他邻居节点完全覆盖和连通的，就算因其担当簇头所进行了复杂的工作耗尽能量而死亡，其监测的位置也不会出现空洞，随着网络监测时间的增长，有效的控制了网络的覆盖率。

在节点随机均匀的布置在监测环境之后，每个节点都开始在传感半径 r 以内的距离干扰其邻居节点并在其路由表中纪录其邻居节点个数。如果数目大于或等于，并且加入剩余能量的考虑，只有在剩余能量大于 10%的初始能量的节点将参与簇头的竞选，竞选的过程采用 LEACH 的随机竞选过程。

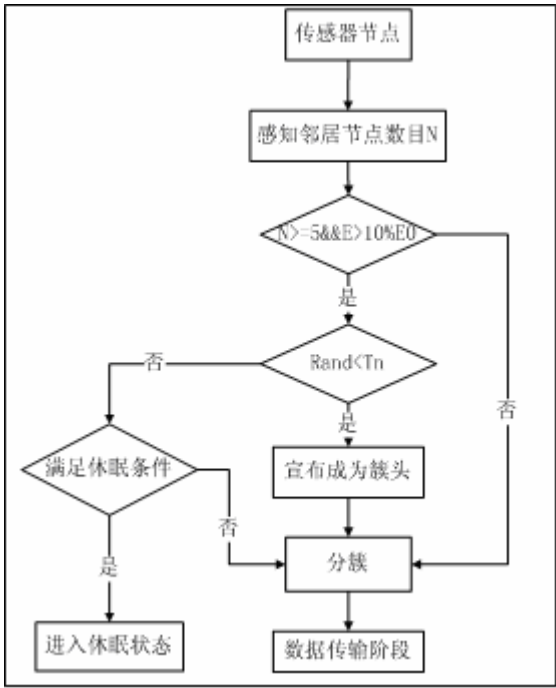


图1 每一轮 LEACH-CC 协议的流程图

选簇阶段完成后，冗余节点依次进行休眠，这里引入一个“退后机制”：当一个节点进入休眠状态时，都会向周围的邻居节点发送一个 SAM（状态广告消息）控制包，其邻居节点收到该控制包都会将自己的邻居数-1，重新统计邻居数，如果在自己的休眠间隙到来时，邻居节点数大于或等于则准备进入休眠状态，其中当有节点死亡，在宣布死亡的同时也向邻居节点发送数据控制包，则它的每个邻居节点都将自己的邻居数-1。准备进入休眠状态的节点再过一个时间 T_W ，如果满足邻居节点数大于或等于就进入休眠状态，如果在此时间内收到邻居节点的 SAM 控制包就激活，统计邻居节点数。这样的休眠模式，避免了邻居节点数大于或等于的节点同时休眠造成网络空洞。当有邻居节点死亡需要唤醒休眠节点的时候，等待唤醒的节点

也将等待一个 T_d 时间间隔，在这个间隔之后如果仍然满足唤醒条件，则唤醒进入工作状态，这样就避免了同时唤醒多个节点，造成冗余节点。剩余的节点就按照距离的远近加入离自己最近的簇头，完成分簇阶段。

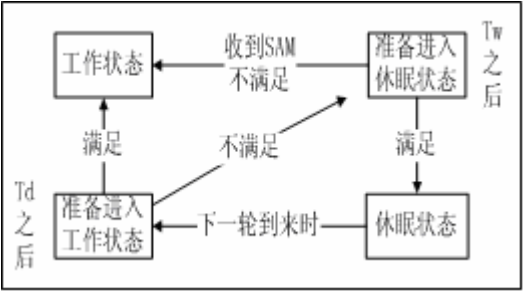


图2 节点的休眠模式

为了防止选举出来的簇头可能会集中在一起，造成网络分布不均，当第一个簇头选举成功后可以在周围广播自己的消息抑制其邻居节点成为簇头。当然可以在网络仿真的过程中加入簇头与簇头之间的距离不得小于某个特殊的值，使其网络分簇更加的均匀更加优化。

这样的分簇方式，既不需要精确的节点的定位，和相互交换其位置信息，通过简单的自动感测，减少了上层控制信息，体现了无线传感器网络快速组网的要求。并且加入了合理的休眠机制，在保证网络的高覆盖率和连通率的前提下，最大限度的节约网络的能量，延长网络的生命周期。

5 性能评价

利用 NS2 网络仿真器^[14]对改进的协议 LEACH-CC 进行仿真分析，并与 LEACH 协议的性能进行比较。LEACH-CC 的指标来评价：（1）10%节点死亡的生命周期；（2）10%节点存活的生命周期；（3）网络覆盖率。本次模拟实验将在 100×100 的区域内随机分布 100~400 个节点，节点感知半径为 10m，采用了文献[2][13]相同的参数和能量模型。

表1 模拟实验参数列表

Parameter	Value
Network size (m^2)	100×100
The number of deployed nodes	100
Sensing range (Rs) (m)	10
$E_{elec}(nJ/bit)$	50
$\varepsilon_{is}(pJ/bit/m^2)$	10
$\varepsilon_{amp}(pJ/bit/m^4)$	0.0013
$E_{fusion}(nJ/bit/signal)$	5
Data packet size (byte)	1500

续表

Parameter	Value
-----------	-------

Broadcast packet size (byte)	50
Packet header size (byte)	50
Round (TDMA frames)	10
Initial energy (J/battery)	0.5

无线对称信道传输能量模型:

当发射数据包为 k 比特, 传输距离为 d 时, 发射端消耗的能量:

$$E_{Tx}(k, d) = \begin{cases} kE_{elec} + k\varepsilon_{fs}d^2 & d < d_0 \\ kE_{elec} + k\varepsilon_{amp}d^4 & d > d_0 \end{cases} \quad (3)$$

其中, d_0 为距离常数, $d_0 = \sqrt{\frac{\varepsilon_{fs}}{\varepsilon_{amp}}}$ 。但一般情况下, 节点的传输距离都要求小于 d_0 。

接收端消耗能量:

$$E_{Rx}(k) = kE_{elec} \quad (4)$$

图 3 是随意的一次分簇结果, 可以明显看出其簇头分布均匀, 其余节点按照离簇头的远近加入离自己最近的簇头, 分簇均匀。图 4 是仿真了 100 个节点, 其中有 10% 个节点死亡时的生命周期比较可以看出, 改进后的 LEACH-CC 协议是 LEACH 协议的 1.37 倍, 图 5 也反应了这个情况, 当几乎全部节点死亡的时候, 改进的 LEACH-CC 的生命周期延长了 0.58 倍, 超过一半。图 6 表示随着网络节点的不断增多, 其改进的协议的生命周期接近 LEACH 协议的两倍。以上说明了改进后的 LEACH-CC 协议增长了 LEACH 的生命周期, 延长网络寿命。接下来分析我们关注的覆盖率的问题。图 7 是对 100 个节点的仿真结果, 随着节点的死亡数目增多对比其网络覆盖率的变化情况。在网络覆盖率达到 70% 的时候, LEACH 协议死亡节点为 61, LEACH-CC 协议死亡节点数为 83。由此可以看出, 改进后的协议随着死亡节点的增加, 对网络的覆盖率有了很大程度的提高。

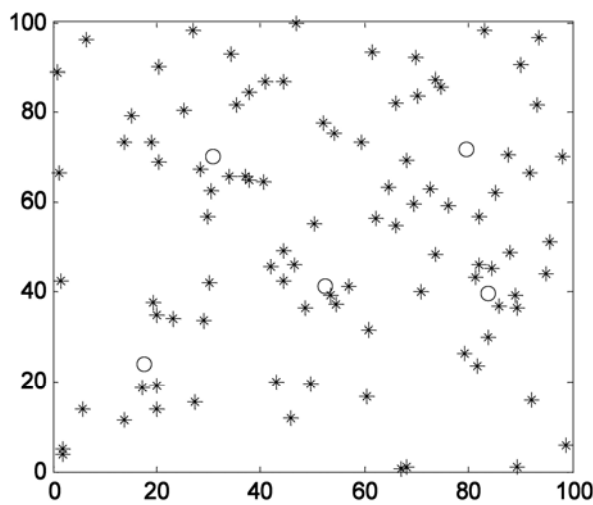


图 3 100 个随机分布节点的分簇情况 (圆形为簇头节点)

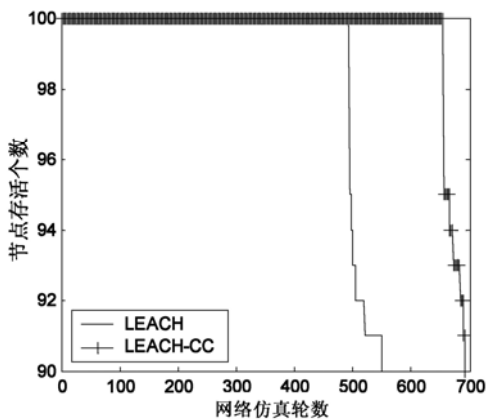


图 4 10%的死亡节点的网络生存周期比较

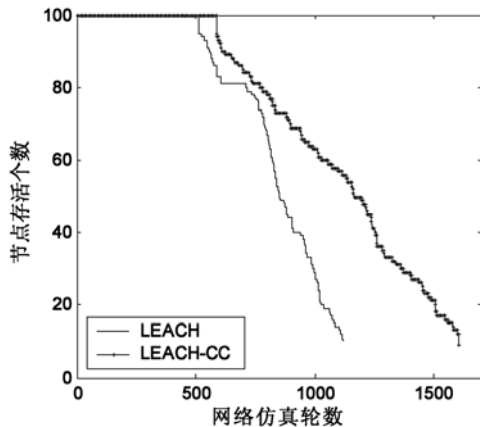


图 5 10%存活节点的网络生存周期比较

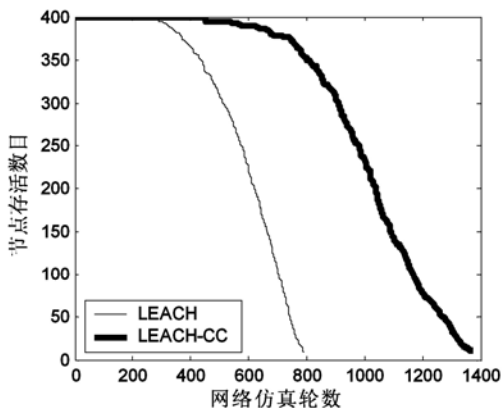


图 6 400 个节点的网络生存周期比较

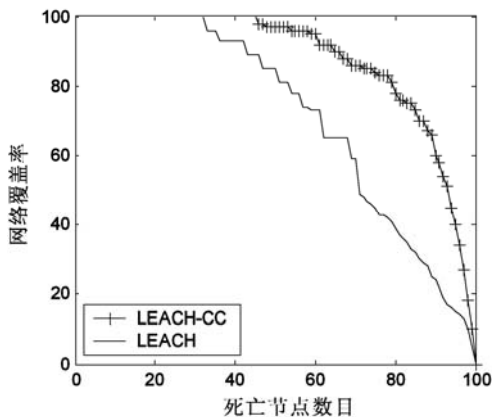


图 7 网络覆盖率比较

6 总结

在这篇论文中，创新的引入了覆盖与连通的性质加入簇头的选举中，从理论上证明了其高覆盖和高连通的可实施性，通过邻居节点数来选举簇头和休眠节点，使得节点的控制消息非常简单适合无线传感器网络。并且通过网络仿真也证实了改进协议 LEACH-CC 的有效的增长了网络的生存时间，也增加了网络的覆盖率，确保了信息的全面有效的采集。今后的工作将注重覆盖和连通的研究，找出网络密度和邻居节点数的关系，分析多种覆盖与连通的模型，以及可以尝试在三维空间中的研究和仿真。

参考文献

- [1] Rentala P, Musunuri R, Gandham S, Saxena U. Survey on sensor networks [R]. Technical Report, UTDCS-33-02, University of Texas at Dallas, 2002.
- [2] Heinzelman W R, Chandrakasan A, Balakrishnan H. Energy-efficient communication protocol for wireless microsensor networks [C]. Proceedings of the 33rd Hawaii International Conference on System Sciences. Maui:IEEE Computer Society, 2000.3005-3014.
- [3] Younis O, Fahmy S. HEED: A Hybrid, Energy-Efficient, Distributed Clustering Approach for Ad-hoc Sensor Networks [C]. In: Proc 13th Joint Conf on IEEE Computer and Communications Societies (INFOCOM), March 2004.

- [4] Stephanie L, Cauligi S R. PEGASIS:Power efficient gathering in sensor information systems [C]. Proceedings of the IEEE Aerospace Conference, Big Sky, Montana, March 2002.
- [5] Ya X, John H, Deborah E. Geography-informed Energy Conservation for Ad Hoc Routing [C]. In Proceedings of the seventh annual international conference on Mobile computing and networking (MobiCom 01), Rome, Italy, July 2001.70–84.
- [6] Fan Y, Gary Z, Songwu L. Energy Efficient Robust Sensing Coverage in Large Sensor Networks[R]. Technical Report, UCLA, 2002.
- [7] Piyush G, P R Kumar. Critical power for asymptotic connectivity[C]. Proceedings of the 37th IEEE conference on decision & control, Tampa, Florida USA, December 1998.1106-1110.
- [8] Sachin A , Mani S. Critical Density Thresholds for Coverage in Wireless Sensor Networks [C]. Proc IEEE Wireless Comm and Networking Conf , 2003.1615-1620.
- [9] Jorge C, Sonia M, Timur K. Coverage Control for Mobile Sensing Networks [J]. IEEE Trans Robotics and Automation, vol 20, issue 2, Apr.2004.243-255.
- [10] Xiaojiang D,Fargo L. Maintaining Differentiated Coverage in Heterogeneous Sensor Networks [J].EURASIP J on Wireless Comm and Networking, vol 5, issue 4, 2005.565-572.
- [11] Yong G, Kui W, Fulu L. Analysis on the redundancy of wireless sensor network [J]. Proe Of The 2nd ACM Int.Conf.on Wireless Sensor Networks and Applications (WSNA2003). San Diego: ACMPress, 2003.108-114.
- [12] Feng X, P R Kumar. The number of neighbors needed for connectivity of wireless Networks [J]. Wireless networks 2004(10):169-181.
- [13] Di Tian, Nicolas D G. Connectivity maintenance and coverage preservation in Wireless sensor network [J].Ad-hoc Network Journal, 2005, 3(6):744-761
- [14] 柯志亨, 程荣祥, 邓德隽. 仿真实验——多媒体和无线网络通信[M]。电子工业出版社。2009。

作者简介:

朱丽娟 (1985—), 女, 汉族, 深圳大学信息工程学院 硕士研究生, 研究方向: 无线传感器网络及能耗控制分析 sy70130102@szu.edu.cn

王敬国 (1984—), 男, 汉族, 深圳大学信息工程学院 硕士研究生, 研究方向: 无线传感器网络及定位算法 sy70130106@szu.edu.cn

林晓辉 (1975—), 男, 汉族, 深圳大学信息工程学院 教授, 研究生导师, 研究方向为无线网络与分布式系统。长期从事移动计算与分布式系统, 无线网络资源管理, 无线网络功率控制与流量调度, 路由算法设计与评估, 优化理论与应用等方面的研究, 主持多个国家及省、市自然科学基金项目、深港创新圈项目等, 并在 IEEE 等国际一流学术期刊与会议上发表论文二十余篇。xhlin@szu.edu.cn

王晖 (1969—), 男, 汉族, 深圳大学信息工程学院 教授, 硕士生导师, 研究方向为信号处理及无线通信。长期从事数字图象处理、智能医学信息处理和数字通信等方面的研究, 主持并完成了多个国家及省、市自然科学基金项目, 是国际国内多个核心期刊审稿人, 并且在国际国内一流学术期刊与会议上发表论文二十余篇。wanghsz@szu.edu.cn

基于量子遗传算法的OFDM自适应调制技术

李兆华 李 飞

(南京邮电大学 通信与信息工程学院, 南京 210003)

摘 要: 针对多用户 OFDM 的自适应调制问题, 利用量子态的并行计算特性和量子纠缠特性, 使用量子遗传算法有效获得自适应调制的分配方案, 解决冲突子载波的分配。仿真结果表明, 此算法在误码率方面具有良好性能并优于静态分配算法和其他动态分配算法。

关键词: 量子遗传算法; OFDM 系统; 频分复用; 自适应调制

The OFDM Adaptive Modulation Based on Quantum Genetic Algorithm

Li Zhao hua Li Fei

(College of Communication & Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing 210003, China)

Abstract: On OFDM adaptive modulation for multi-user, use quantum genetic algorithm to get the distribution of adaptive modulation. Taking advantage of the properties of quantum mechanics such as quantum parallelism and entanglement, can allocate the conflict subcarriers effectively. The simulation results shows that this algorithm has a good BER performance, which is superior to the static and other dynamic allocation algorithm.

Keywords: quantum genetic algorithm; OFDM; frequency division multiplexing access; adaptive modulation

1 引言

正交频分复用技术 (OFDM) 能很好的抑制符号间干扰, 已成为下一代移动通信系统物理层的关键技术。在 OFDM 多用户系统中, 对某个用户信道质量很差的子载波, 对另一个用户来说信道质量可能会非常好。自适应调制技术可以根据信道质量情况决定调制方式, 从而充分利用信道特性, 提高信道的频谱利用率。在保证系统的误码率和传输速率的前提下, 由于不同子载波的信道衰落不同, 从而需要的发送功率也不同, 因此需要进行适当的功率分配, 这就是 OFDM 中的最优化功率分配问题。

在单用户情况下, OFDM 系统的所有子载波都分配给单个用户使用, 因此不存在子载波多用户分配问题。在文献[1]中, 灌水法已经被证明是在满足系统误码率要求下使系统发射功率最小的最优算法。而在多用户情况下, 由于多个用户不能共享同一子载波, 对一个子载波的分配将严重影响妨碍其他用户对此子载波的使用, 因此灌水法不再成为最优的解决方法。

频分复用技术（FDMA）是一种静态频带分配方法，不考虑每个用户的信道信息，而是为每个用户分配固定的子载波，算法简单，但多用户分集增益没有得到充分利用。

为了充分利用信道信息，可以使用动态算法进行多用户子载波比特分配。文献[2]提出使用反复灌水法来解决子载波和比特分配问题，避免复杂的计算，效果较好。而文献[3]则提出使用拉格朗日松弛算法进行子载波和比特的动态分配，也可以达到文献[2]中提出的算法性能，但算法非常复杂。量子遗传算法（QGA）^[4]在解决组合优化问题中表现出了良好性能，本文尝试使用量子遗传算法来解决多用户 OFDM 系统的自适应调制问题，仿真结果表明，其性能优于文献[2]和文献[3]中的算法，且计算复杂度较低。

2 OFDM系统模型

假设 OFDM 系统有 K 个用户，第 k 个用户的数据速率为 R_k 比特/OFDM 符号。系统中共有 N 个子载波。则可设 $P_{k,n}$ 为用户 k 在子载波 n 上的功率， P_{total} 为每个 OFDM 符号需要的总的功率， $\alpha_{k,n}$ 为用户 k 在子载波 n 上的信道衰落， $b_{k,n}$ 为用户 k 在子载波 n 上的比特数，则优化目标为最小发射功率的动态子载波，比特分配问题可如下描述：

$$\arg \min P_{total} = \sum_{n=1}^N \sum_{k=1}^K \frac{P_{\min}(b_{k,n})}{\alpha_{k,n}^2} \quad (1)$$

$$\text{s t} \quad \sum_{n=1}^N b_{k,n} = R_k \quad \forall k, \quad (2)$$

$$BER_n \leq BER_{target} \quad (3)$$

$$P_{\min}(b_{k,n}) = N_0 / 3 \left[Q^{-1} \left(\frac{BER_n}{4} \right) \right]^2 (2^{b_{k,n}} - 1) \quad (4)$$

其中式(2)表示每个用户传送的比特约束，式(3)规定了每个用户的误比特率上限，式(4)表示了信道增益为 1 时，为保证用户 k 在子载波 n 上每个调制符号可靠传输 $b_{k,n}$ 比特所需的最小功率。

3 基于量子遗传算法的多用户自适应调制方案

量子遗传算法（Quantum Genetic Algorithm, QGA）是量子计算与遗传算法相结合的产物。该算法用量子比特编码来表示染色体，使得一条染色体可以表达多个态的叠加，并利用量子旋转门或量子非门实现染色体的更新操作，从而实现目标优化求解。与经典遗传算法相比，具有种群规模小，寻优能力强，收敛速度快,能高速并行等优点而受到极大关注。

3.1 量子比特编码

量子遗传算法中，使用量子比特（qubit）来表示一个基因，一个量子比特可能处于 $|0\rangle$ 状态或 $|1\rangle$ 状态，或者 $|0\rangle$ 和 $|1\rangle$ 之间的中间态。一个包含 m 个 qubit 的量子染色体可描述为：

$$\begin{bmatrix} \alpha_1 | \alpha_2 | \dots | \alpha_k \\ \beta_1 | \beta_2 | \dots | \beta_k \end{bmatrix}$$

$$|\alpha_i|^2 + |\beta_i|^2 = 1$$

其中 $|\alpha_i|^2$ 和 $|\beta_i|^2$ 分别表示量子比特处于 0 和 1 状态的概率。其中 $i=1,2,\dots,k$

3.2 量子遗传算法步骤描述

量子遗传算法步骤与经典遗传算法类似，但其种群由量子染色体构成，设 t 代的量子染色体种群为 $Q(t) = \{q_1^t, q_2^t, \dots, q_n^t\}$ ，其中 n 为种群大小， t 为遗传代数， m 为量子染色体长度， q_j^t 为定义如下的量子染色体，此量子染色体的长度定义为 m ：

$$q_j^t = \begin{bmatrix} \alpha_1^t | \alpha_2^t | \dots | \alpha_m^t \\ \beta_1^t | \beta_2^t | \dots | \beta_m^t \end{bmatrix}, j=1,2,\dots,n$$

则 QGA 步骤描述如下：

- ① 初始化种群 $Q(t)$ ，即生成第 0 代种群 $Q(0)$ ；
- ② 由 $Q(t)$ 量子坍塌生成个体 $P(t)$ ；
- ③ 评价 $P(t)$ 中每个个体的适应度；
- ④ 选择并保存适应度最高的个体及其适应度值；
- ⑤ 迭代终止条件判断：当满足终止条件时，输出当前最优个体，否则继续进行下一步；
- ⑥ 进行量子交叉操作，更新 $Q(t)$ ；
- ⑦ 进行量子变异操作，使用量子旋转门变异操作更新 $Q(t)$ ，得到下一代种群 $Q(t+1)$ ；
- ⑧ 转到②，继续进行遗传操作；

与经典遗传算法相比，量子遗传算法不同处在于：①中，初始化时，初始种群的染色体所有基因 α_i^0, β_i^0 全部设为 $1/2^{1/2}$ ，即表示量子染色体坍塌为 0, 1 状态的概率相同。②中，通过量子坍塌以后所生成的 $P(t) = \{x_1^t, x_2^t, \dots, x_n^t\}$ ，其中 x_i^t 为第 t 代种群的第 i 个解，具体值为长度为 k 的二进制串，生成方法为：随机产生一个属于 $[0,1]$ 的数，若它大于 $|\alpha_i^t|^2$ ，则 x_i^t 取 1，否则取 0。⑥中，量子交叉的具体方法为：以一定的概率从 P_m 中随机选择个体，对于选中的个体随机确定一个交叉位，最后将该个体的交叉位的几率幅位置对调，即将 $|0\rangle$ 状态调整为 $|1\rangle$ 状态或相反。这样做可以阻止未成熟收敛和加强算法的全局搜索能力。

3.3 量子变异操作

根据遗传算法对变异操作的定义，为了充分利用量子计算高度并行性的特点，可以采用量子旋转门策略对量子染色体进行变异，其定义如下：

$$\begin{pmatrix} \alpha_i' \\ \beta_i' \end{pmatrix} = \begin{pmatrix} \cos(\theta_i) & -\sin(\theta_i) \\ \sin(\theta_i) & \cos(\theta_i) \end{pmatrix} \begin{pmatrix} \alpha_i \\ \beta_i \end{pmatrix}$$

其中， (α_i, β_i) 为量子染色体中第 i 个 qubit 基因， θ_i 为旋转角，由下式决定：

$\theta_i = s(\alpha_i, \beta_i) \Delta \theta$ 。其中 $s(\alpha_i, \beta_i)$ 和 $\Delta \theta$ 分别代表旋转的方向和角度。其值由表 1 所示策略决定：

表 1 量子变异操作旋转角改变策略

x_i	b_i	$f(x)>f(b)$	$\Delta \theta_i$	$s(\alpha_i, \beta_i)$			
				$\alpha_i \cdot \beta_i > 0$	$\alpha_i \cdot \beta_i < 0$	$\alpha_i = 0$	$\beta_i = 0$
0	0	False	0	-	-	-	-
0	0	True	0	-	-	-	-
0	1	False	0.02π	+1	-1	0	± 1
0	1	True	0.02π	-1	+1	± 1	0
1	0	False	0.02π	-1	+1	± 1	0
1	0	True	0.02π	+1	-1	0	± 1
1	1	False	0	-	-	-	-
1	1	True	0	-	-	-	-

其中 x_i 为当前染色体 i 位， b_i 为最优染色体 i 位。 $f(x)$, $f(b)$ 为各自的自适应度函数。

3.4 使用量子遗传算法进行自适应调制

本文提出的基于量子遗传算法的 OFDM 系统多用户自适应调制方案中，首先对各用户独立进行贪婪灌水算法，对于不冲突子载波，进行固定分配，然后对于出现冲突的子载波，使用量子遗传算法进行子载波分配，具体参数设置如下：

① 染色体编码方案：种群大小设为需要分配的子载波数目 N_c ，每个染色体代表一个待分配的子载波，染色体长度 m 根据系统用户数 k 确定($2^m \geq k$)。

② 种群初始化及量子坍塌方法均与前面介绍相同，量子坍塌后生成的每个个体即代表一种子载波分配方案，如下表所示：

表 2 子载波分配方案

子载波 1	子载波 2	...	子载波 N_c
-------	-------	-----	-----------

③ 适应度函数评价：针对每个个体对应的子载波分配方案，结合固定分配方案，使用贪婪灌水法针对每个用户进行比特分配，计算所有用户需要的总功率，作为评价个体的适应度准则。总功率越小，则适应度越高。

④ 量子变异操作：除了以上提到的变异策略外，为了有效利用当前代信息，使用以下变异操作：首先由当前最优个体推出一个指导染色体，然后在它周围随机散布量子染色体最为下一代的量子种群，用公式表示

$$Q_g(t) = \alpha \cdot P_c(t) + (1 - \alpha) \cdot (1 - P(t))$$
$$Q(t+1) = Q(t) + b \cdot \text{normrnd}(0,1)$$

其中 $\alpha \in [0.1, 0.5]$, $b \in [0.05, 0.15]$

⑤ 终止条件：设定遗传操作代数 G ，当算法迭代次数达到 G 时终止。

4 仿真实验

针对多用户的 OFDM 系统，假设系统信道是准高斯信道，随机生成服从复高斯分布的各子载波信道响应，可以采用的调制方式为 BPSK, QPSK, 16-QAM, 64-QAM, 256-QAM。设置 OFDM 系统子载波数为 64 个，分配给 4 个用户，每个用户在一个 OFDM 符号中需要传输的比特数为 20，分别使用以下算法进行仿真：

OFDM-FDMA 算法：将 64 个子载波平均分配给 4 个用户，每个用户使用固定的 16 个子载波进行调制。

文献[2]中算法：首先对各用户进行贪婪灌水算法，然后对出现冲突的子载波进行动态分配，从总功率最大的子载波开始，将此子载波单独分配个某用户，保证分配后使系统的总功率增加最少。然后其他用户重新进行贪婪灌水算法比特分配，直到所有的子载波分配完毕。

文献[3]中算法：首先放宽子载波比特分配的条件，假定可以为每个子载波分配任意比特（包括小数），然后引入非线性规划方法，使用拉格朗日乘子法进行求解，最后将求得结果调整为符合要求的比特数目。

本文中的量子遗传算法：对各用户单独使用贪婪灌水算法进行比特分配，然后对冲突子载波使用量子遗传算法进行分配，种群数目设为 20，遗传迭代 100 次。

4.1 算法收敛速度比较

为比较量子遗传算法应用于子载波调制的收敛速度，使用相同的参数，进行了经典遗传算法的仿真，如图 1 所示：

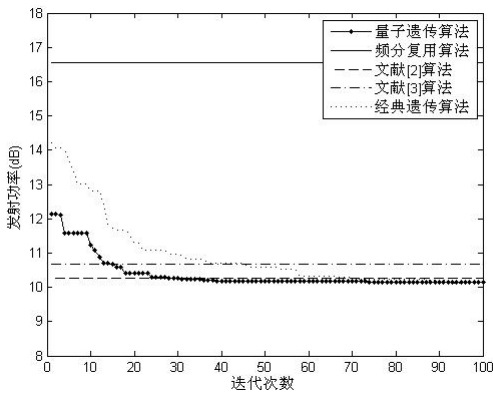


图 1 量子遗传算法的收敛速度

由上图可以看出，量子遗传算法在一开始即优于频分复用（FDMA）算法，在 30 代以内可以收敛到优于文献[2]的结果，同时量子遗传算法的收敛速度也明显优于经典遗传算法。

4.2 仿真结果

对于以上四种算法，进行蒙特卡洛仿真 100 次，仿真结果如下图所示。

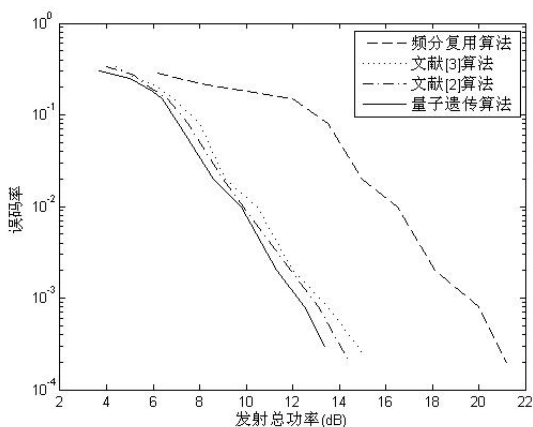


图2 不同算法的发射功率—误码率图

由图2可以看出,量子遗传算法大大优于FDMA静态分配算法,也稍优于文献[2]和文献[3]提出的动态分配算法。

5 结束语

本文提出了一种基于量子遗传算法进行OFDM自适应调制的方案,利用量子计算的特点如量子并行计算,量子纠缠等特性,使量子遗传算法具有比经典遗传算法更强的并行处理能力和更快的收敛速度,仿真结果表明,本文提出的算法具有比传统的载波分配算法更好的性能。

参 考 文 献

- [1] Cover TM, Thomas TA. Elements of Information Theory [M]. New York: Wiley, 1991.
- [2] Guo-dong Zhang. Subcarrier and bit allocation for real-time services in multiuser OFDM systems [C].IEEE International Conference on Communications. Paris, France:2004 Vol 5. 2985-2989.
- [3] Cheong Yui Wong, Roger S. Cheng, Khaled B L and Ross D M. Multiuser OFDM with Adaptive Subcarrier, Bit, and Power Allocation [J]. IEEE JASA,1999,17(10):1747-1758
- [4] 杨淑媛, 刘芳, 焦李成.一种基于量子染色体的遗传算法[J].西安电子科技大学学报, 2004, 31(1):76-81

基于蚁群算法的OFDMA系统自适应资源分配算法

张 昕¹ 叶 梧¹ 冯穗力¹ 庄宏成²

(1.华南理工大学 电子与信息学院, 广东 广州 510640;

2.华为技术有限公司, 广东 深圳 518129)

摘 要: 提出了一种基于蚁群算法的 OFDMA 系统自适应资源分配方案, 将原问题建模成一个寻找最优路径的问题, 并运用蚁群算法使蚂蚁在信息素及启发式信息的引导下构建最优的路径, 从而得到原问题的解。由于蚁群算法所具有信息正反馈以及启发式搜索等特点, 使得算法可以通过较少的迭代得到一个较好的分配方案, 与最优值相比仅相差 1dB, 而与静态子载波分配算法相比性能可以提高 1~2dB。

关键词: 资源分配; OFDMA; 蚁群算法

Adaptive Resource Allocation for OFDMA System Based on Ant Colony Algorithm

Zhang Xin¹ Ye Wu¹ Feng Sui Li¹ Zhuang Hong Cheng²

(1. School of Electronic and Information Engineering, South China University of Technology, GuangZhou 510640, China; 2. Huawei Technologies Co. Ltd, ShenZhen 518129, China)

Abstract: This paper studied the problem of adaptive resource allocation in OFDMA system. An ant colony algorithm based adaptive resource allocation was proposed in this paper. The problem was modeled as finding a minimum cost path in a graph. The ants' solution construction was guided by pheromone trail and heuristic information. Due to the unique heuristic searching mechanism of the ant colony algorithm, the proposed algorithm is guaranteed to converge quickly to an approximately optimal solution. Simulation results show that the performance of the proposed algorithm is close to the optimal solution, and gains 1-2dB in comparison with the fixed allocation scheme.

Keywords: resource allocation; orthogonal frequency division multiple access(OFDMA) ; ant colony algorithm

1 引言

OFDMA 是以 OFDM 调制技术为基础, 通过给每个用户分配一系列的子载波来实现多用户接入的多址技术。OFDMA 系统可以根据用户的信道条件为每个用户安排合适的子载波进行传输, 实现动态的资源分配, 从而有效地利用系统的功率和频谱资源, 提高系统的性能, 并且能够有效对抗由于多径效应引起的窄带干扰。因此 OFDMA 技术被广泛应用于各种通信系统, 成为宽带无线通信系统的关键技术, 在 802.16 的无线城域网 (WMAN)、3GPP-LTE 中都

将 OFDMA 作为其物理层的技术。

OFDMA 系统的资源分配问题就是研究如何为不同的用户分配子载波、比特以及功率。根据优化目标的不同，可以分为下面两种形式：一种是基于总发射功率最小化的 MA(Margin Adaptive)问题^[1,2]，它的优化目标是在用户数据速率固定的条件下，使系统的发射功率最小化；另一种是基于吞吐量最大化的 RA(Rate Adaptive)问题^[3,4]，它的优化目标是在总发射功率固定的条件下，使系统的容量最大化。

文献[1]最早对 MA 问题进行了研究，作者首先建立了在用户速率以及误比特率等约束条件下最小化基站总发送功率的非线性模型，并利用拉格朗日松弛算法对求得了这个非线性问题的次优解，但是这种算法需要同时搜索两个拉格朗日参数，其计算复杂度非常高。文献[5]将优化问题分成两个步骤，第一步首先确定各用户需要的子载波数目，第二步将具体的子载波以及相应的功率分配给用户。该算法具有较低的复杂度，并且能获得与文献[1]性能相近的解。文献[2]将 MA 问题建转化成为一个线性规划问题，并利用整数规划的方法给出了问题的最优解。但是整数规划是一个非多项式时间算法，其计算复杂度随变量以及约束条件的数量呈指数增长。文献[6]提出分步式算法，首先采用贪婪算法初步为每个用户分配子载波及比特，然后通过在不同用户间进行子载波及比特的互换来进一步降低系统的功率，但是该算法需要大量的子载波与比特的互换才能收敛到文献[2]的近似解。

本文提出一种基于蚁群算法的自适应资源分配算法，其基本思想是首先将 MA 问题建模成一个二部图 (bipartite graph)，图中的多个边的集合组成一条可行路径，对应 MA 问题的一种分配方案，而求 MA 问题的最优解就对应着寻求图中的一条最优路径。然后将蚂蚁随机地放置在图中的用户集上，让蚂蚁在信息素及启发式信息的引导下构建最优的路径。由于蚁群算法的所特有的信息正反馈和启发式搜索等特征，使得算法能够通过有限次的迭代快速收敛到一个近似最优解。

2 系统模型与问题描述

2.1 系统模型

OFDMA 的系统模型如图 1 所示：系统中子载波数为 N ，总用户数为 K ，每个 OFDM 符号周期内的用户比特数为 R_k ($k=1,2,\dots,K$)，假设基站能够通过反馈信道获得用户端的信道状态信息 (CSI) 的值，并根据该值动态地为每个用户分配子载波、比特以及功率。 $\rho_{k,n}$ 表示子载波的占用情况， $\rho_{k,n}=1$ 表示子载波 n 被用户 k 占用，否则 $\rho_{k,n}=0$ 。 $h_{k,n}$ 表示用户 k 在

子载波 n 上的信道增益， $c_{k,n}$ 与 $p_{k,n}$ 分别表示为用户 k 在子载波 n 上分配的比特数及功率，在一定的误码率 P_e 条件下， $c_{k,n}$ 与 $p_{k,n}$ 的关系可以表示为：

$$p_{k,n} = \frac{f_k(c_{k,n})}{|h_{k,n}|^2} \quad (1)$$

其中 $f(\cdot)$ 表示信道增益为 1 时，在一定的误码率条件下，接收 c 个比特的数据所需要的接收功率，在使用 M-QAM 调制方式的情况下， $f_k(c)$ 可以表示为^[8]：

$$f_k(c) = \frac{N_0}{3} [Q^{-1}(\frac{p_e}{4})]^2 (2^c - 1) \quad (2)$$

式中 p_e 表示误码率, N_0 表示噪声功率谱密度, $Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-t^2/2} dt$ 。

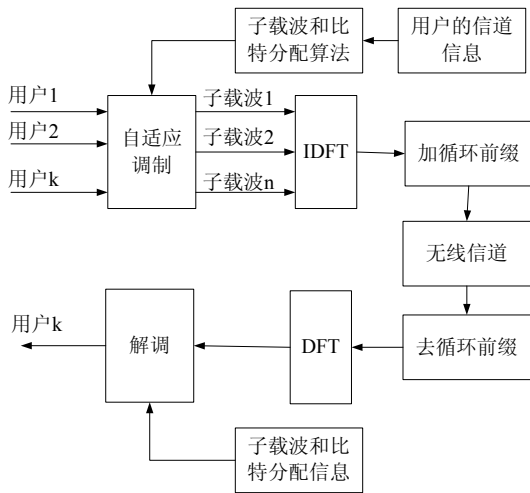


图1 OFDMA 的系统模型

2.2 问题描述

MA 问题就是在满足用户速率要求的情况下,通过子载波分配和比特加载,使系统的总发射功率最小化,这个问题可以用数学表述为:

$$P_T = \min \sum_{n=1}^N \sum_{k=1}^K \frac{\rho_{k,n}}{h_{k,n}^2} f_k(c_{k,n}) \quad (3a)$$

$$s.t. \quad R_k = \sum_{n=1}^N \rho_{k,n} \cdot c_{k,n} \quad \forall k \in \{1,2,\dots,K\} \quad (3b)$$

$$\sum_{k=1}^K \rho_{k,n} = 1 \quad \forall n \in \{1,2,\dots,N\} \quad (3c)$$

式中 P_T 表示系统的总功率,约束条件式(3b)为系统中各个用户的速率要求,约束条件(3c)保证每个子载波只能被一个用户占用。由于 $f_k(c)$ 是非线性的,因此 MA 问题是一个非线性规划问题。

3 基于整数规划的自适应资源分配算法

根据文献[2,6],上述的问题可以转化为一个整数规划问题进行求解。我们首先定义一个变量 $x_{k,n,c} \in \{0,1\}$, $x_{k,n,c} = 1$ 表示用户 k 占用第 n 个子载波,并且在该子载波上分配的比特数为 c ,否则 $x_{k,n,c} = 0$ 。另外定义一个变量 $p_{k,n,c}$ 表示用户 k 在子载波 n 上发送 c 比特信息所需要的发送功率,则:

$$p_{k,n,c} = \frac{f_k(c_{k,n})}{|h_{k,n}|^2} \quad (4)$$

对任一用户 k ，我们都可以通过公式 (2) 计算得到 $f_k(c_{k,n}) \in \{f_k(1), \dots, f_k(C)\}$ ，因此原 MA 问题可以转化为关于变量 $x_{k,n,c}$ 的整数规划问题：

$$\begin{aligned} \min \quad & \sum_{k=1}^K \sum_{n=1}^N \sum_{c=1}^M p_{k,n,c} \cdot x_{k,n,c} \\ \text{s.t.} \quad & R_k = \sum_{n=1}^N \sum_{c=1}^M c \cdot x_{k,n,c} \\ & 0 \leq \sum_{k=1}^K \sum_{c=1}^M x_{k,n,c} \leq 1 \\ & x_{k,n,c} \in \{0,1\} \end{aligned} \quad (5)$$

上述整数规划问题变量的数量为 $K \times M \times C$ 个，约束条件的数量为 $N+K$ 个，可以运用整数规划的方法进行求解。然而整数规划是非多项式时间算法 (Nondeterministic Polynomial Time)，其计算复杂度随变量以及约束条件的数量呈指数增长。因此从实用的角度出发，我们需要设计一种低复杂度，性能接近最优解的算法。

4 基于蚁群算法的自适应资源分配算法

4.1 蚁群算法描述

蚁群优化算法^[7]是一种分布式智能算法，其基本思想是模仿蚂蚁依赖信息素进行通信，从而互相协作，完成从蚁穴到食物源寻找最短路径的行为。它是一种随机的通用试探方法，具有信息正反馈和启发式搜索的特征，最早是用来解决旅行商问题，后来经过不断的发展，在求解各种不同的组合优化问题中获得了广泛的应用。为了能将蚁群算法应用到 MA 问题中，可以把 MA 问题表示为一个二部图 $G = (V, U, E)$ ，如图 2 所示：

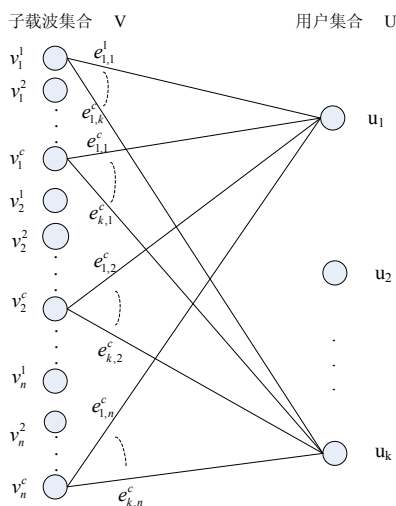


图 2 二部图

图中集合 V 表示 $N \times C$ 个节点的子载波集合, 对于每个子载波 n , $v_n^c \in \{v_n^1, \dots, v_n^C\}$, v_n^c 表示子载波 n 上发送的比特数为 c ; U 表示 K 个节点的用户集合, $u \in \{u_1, \dots, u_K\}$; E 是连接用户节点与子载波节点的边的集合, $E = \{e_{k,n}^c | k=1, \dots, K; n=1, \dots, N; c=1, \dots, C\}$, 如果子载波 n 分配给用户 k , 并且在子载波 n 上发送的比特数为 c , 则 v_n^c 与 u_k 之间有边 $e_{k,n}^c$ 相连。图的多个边组成一条可行路径, 对应 MA 问题中的一种分配方案, 求 MA 问题的最优解就对应着寻求图中的一条最优路径。蚁群算法的基本思想就是将蚂蚁随机的放置在图中的用户集 U 上, 让蚂蚁根据信息素及启发式信息寻找最优的路径, 从而构建问题的解。

4.2 蚁群算法主要步骤

步骤 1. 初始化: 将 m 只蚂蚁随机放置在用户集 U 上, 并设置子载波集 V 与用户集 U 之间的信息素 $\tau_{k,n}^c$ 以及启发式信息 $\eta_{k,n}^c$ 的初始分布。

信息素 $\tau_{k,n}^c$ 表示将子载波 n 分配给用户 k , 并且在子载波 n 上发送 c 比特数据的期望度, 在本文中定义为: $\tau_{k,n}^c = \tau_0 + \Delta\tau$, 其中 τ_0 为初始时刻各条边上的信息素, 为一个较小的值; $\Delta\tau$ 为先验信息。

$$\Delta\tau = \begin{cases} 1, & \left\lfloor \frac{\sum_{k=1}^K R_k}{N} \right\rfloor - 1 \leq c \leq \left\lfloor \frac{\sum_{k=1}^K R_k}{N} \right\rfloor + 1 \\ 0, & \text{否则} \end{cases} \quad (6)$$

启发式信息 $\eta_{k,n}^c$ 使得蚂蚁偏向于选择那些消耗功率比较小的子载波, 在本文中定义为:

$$\eta_{k,n}^c = \frac{2^c}{p_{k,n,c}} \quad (7)$$

步骤 2. 每只蚂蚁 k 按照一定的概率为用户选择子载波, 直到满足用户的速率需求;

$$p_{k,n}^c = \frac{[\tau_{k,n}^c]^\alpha [\eta_{k,n}^c]^\beta}{\sum_{j \notin Tabu_k} [\tau_{k,n}^c]^\alpha [\eta_{k,n}^c]^\beta} \quad (8)$$

上式的 α 与 β 为两个参数, 分别反映了蚂蚁在移动过程中所积累的信息素和启发式信息在路径选择中的相对重要性, $Tabu_k$ 为第 k 只蚂蚁的禁忌表, 即已经被分配给第 k 只蚂蚁的子载波集合。

步骤 3. 蚂蚁移动到下一个未分配资源的用户上, 重复步骤 2;

步骤 4. 所有蚂蚁都遍历完所有用户节点之后, 就建立了 m 个解。计算每个解的功率消耗, 选择最优解进行信息素的更新, 更新规则为:

$$\tau_{k,n}^c = (1 - \rho)\tau_{k,n}^c + \Delta\tau_{k,n}^c \quad (9)$$

其中 ρ 为信息素挥发参数, $\Delta\tau_{k,n}^c$ 定义如下:

$$\Delta\tau_{k,n}^c = \begin{cases} \frac{Q}{L_{gb}}, & \text{如果边 } e_{k,n}^c \text{ 属于路径 } L_{gb} \\ 0, & \text{否则} \end{cases} \quad (10)$$

其中 L_{gb} 表示到目前为止找到的全局最优解。Q 为算法循环一次释放在路径上的信息素总量，用于强化正反馈性能，有助于算法收敛。

5 算法仿真

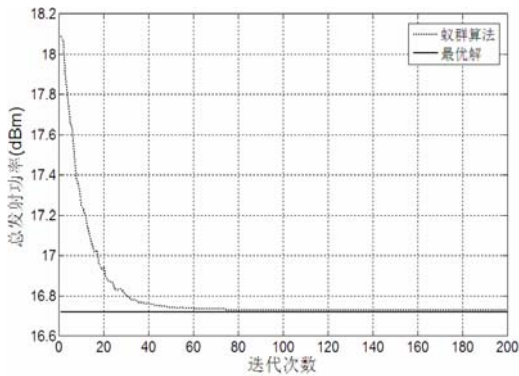
5.1 仿真参数

系统参数：系统带宽为 1MHz，子载波数为 64，噪声功率谱密度为 -80dBW/Hz，信道模型采用 6 径瑞利衰落信道，调制阶数 $c \in \{0,1,2,\dots,6\}$ ，每个用户在一个 OFDM 符号中需要传输的比特数为 20bit，误码率要求为 10^{-4} ，系统的用户数 k 为 2~16。

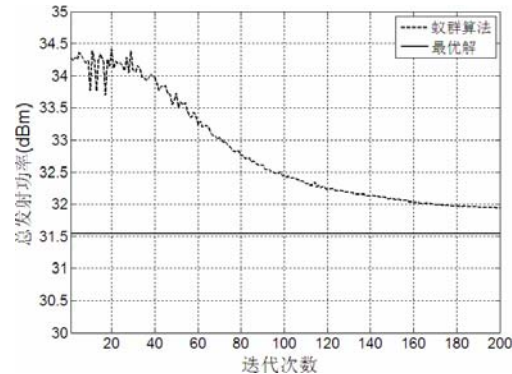
蚁群算法参数： $\alpha = 1; \beta = 2; \rho = 0.1; Q = 50; m = 30; \tau_0 = 0.01$

5.2 收敛性分析

在不同用户数的情况下对蚁群算法分别仿真 10 次后取平均值，可以得到蚁群算法在不同用户数下的收敛性曲线如图 3 所示，由图可见，在用户数较少的情况下，算法的收敛速度较快，当用户数为 2 时，算法在 80 次迭代后收敛，与最优解的差距小于 0.1dB；而随着用户数的增加，计算的复杂度也随之增加，收敛的速度变慢，当用户数为 16 时，算法在 180 次迭代后收敛，与最优值的差距为 0.5dB。



(a) 2 个用户情况下的收敛性能



(b) 16 个用户情况下的收敛性能

图 3

5.3 算法性能分析

为反映算法的性能，将本文所提算法与基于整数规划的最优算法以及静态子载波分配算法（OFDM-FDMA）进行比较，可以得到在不同用户数下的各种算法所需发送功率的曲线如图 4 所示。其中 OFDM-FDMA 算法采用固定的子载波的分配方案,根据各个用户的速率需求按比例预先分配子载波。从图中可以看出，由于采用了动态分配的方法，充分利用了多用户分集增益，因此本文所提算法的性能比静态子载波分配算法相比提高了 1~2dB，而与基于整数规划的最优算法相比仅相差 1dB。

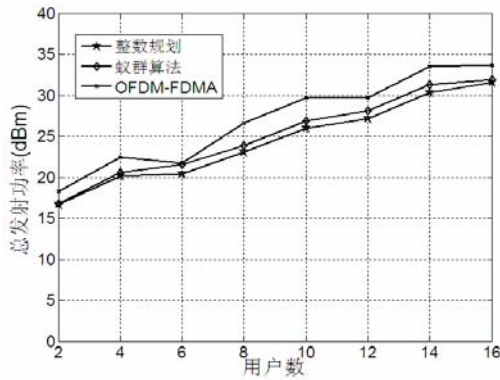


图 4 不同算法下的用户数与总发送功率的关系

6 结论

本文提出了一种基于蚁群算法的 OFDMA 系统自适应资源分配方案，首先将 MA 问题建模成一个二部图 (bipartite graph)，而对原问题的求解等效为在图中寻找一条最优路径的问题，然后利用蚁群算法使蚂蚁在信息素及启发式信息的引导下构建最优的路径。由于蚁群算法具有信息正反馈以及启发式搜索等特点，使得算法可以通过较少的迭代收敛到一个较好的分配方案。仿真结果表明，所提算法的性能与最优值相比仅相差 1dB,而与静态子载波分配算法相比性能可以提高 1~2dB。

参 考 文 献

[1] C. Y. Wong, R. S. Cheng, K. B. Letaief, and R. D. Murch, Multiuser OFDM with adaptive subcarrier, bit, and power allocation[J], IEEE J.Select.Areas Commun., vol.17, no. 10, pp.1747-1758, Oct.1999.

[2] I. Kim, H. L. Lee, B. Kim, and Y. Lee, On the use of linear programming for dynamic subchannel and bit allocation in multiuser OFDM[C], IEEE Global Telecommunications Conference, vol.6, pp. 3648-3652, Nov. 2001.

[3] Z. Shen, J. G. Andrews, B. L. Evans, Adaptive resource allocation in multiuser OFDM systems with proportional rate constraints[J], IEEE Trans. Wireless Commun., Vol.4 No. 6, pp. 2726-2737, Nov 2005.

[4] I. C. Wong, B. L. Evans, Optimal downlink OFDMA resource allocation with linear complexity to maximize ergodic rates[J], IEEE Trans. Wireless Communications, vol.7 no.3, pp: 962-971.Mar 2008

- [5] D. Kivanc, G. Li, and H. Liu, Computationally efficient bandwidth allocation and power control for OFDMA[J], IEEE Trans. Wireless Commun., vol. 2, no. 6, pp. 1150-1158, Nov. 2003.
- [6] M. Ergen, S. Coleri, and P. Varaiya, QoS aware adaptive resource allocation techniques for fair scheduling in OFDMA based broadband wireless access systems[J], IEEE Trans. Broadcast., vol. 49, no. 4, pp. 362-370, Dec. 2003.
- [7] M. Dorigo, V. Maniezzo, and A. Colomi, The ant system: Optimization by a colony of cooperating agents[J], IEEE Trans.Syst.Man Cybern.B, vol.26, pp.29-41, Feb.1996.
- [8] J. G. Proakis, Digital Communications, 3rd ed[M], Nes York: Mcgraw Hill, 1995

作者简介:

张昕(1971-),男,福建人,博士研究生,研究方向:宽带无线通信中的资源分配;通信地址:广东省 广州市 天河区 五山路 华南理工大学 逸夫科学馆 306 室; 邮编 510641; 电子邮箱: xin.zb@mail.scut.edu.cn

叶梧(1938-),男,广东人,教授,博士生导师,研究方向:宽带无线通信;

冯穗力(1955-),男,广东人,教授,博士生导师,研究方向:信息网络理论与技术、宽带无线通信;

庄宏成(1973-),男,博士,华为技术有限公司。

用于多小区OFDMA系统的基于粒子群的 分布式功率分配方法

孙文文 贺志强 牛 凯 田宝玉

(北京邮电大学 信息与通信学院, 北京 100876)

摘 要: 从提高系统频谱效率的角度出发, 提出了一种应用在多小区正交频分多址蜂窝系统下行链路中, 基于粒子群优化方法执行分布式功率分配的方法。为了减轻共信道干扰, 基站在每个小区运用粒子群算法执行功率分配; 且在这个过程中, 每个基站独立工作并不需要与邻近小区交换任何信息。仿真结果显示本文提出的方法相对于两种传统方法, 极大地提高了系统的吞吐量。

关键词: 通信技术; 分布式功率分配; PSO; OFDMA

Particle Swarm Optimization for Distributed Power Allocation in Multi-cell OFDMA Systems

Sun Wen wen He Zhi qiang Niu Kai Tian Bao yu

(School of Information and Communication Engineering , Beijing University of Posts and
Telecommunications ,Beijing ,100876 ,China)

Abstract: In order to increase spectral efficiency, a novel method based on Particle Swarm Optimization (PSO) was proposed to perform distributed power allocation for downlink multi-cell Orthogonal Frequency Division Multiplexing Access (OFDMA) networks. In order to mitigate co-channel interference, base station implemented power allocation by using PSO algorithm in each cell which is partitioned into three sectors. Moreover, each base station could work independently and did not have to change any information with neighboring cells during the process. Simulation results showed that the proposed scheme significantly increases the system's throughput compared with two other algorithms.

Keywords: communication technology; distributed; power allocation; PSO; OFDMA

1 引言(introduction)

正交频分多址接入技术(Orthogonal Frequency Division Multiplexing Access, 简称 OFDMA)已经成为无线网络中一种重要的接入技术, 在 OFDMA 蜂窝系统中的资源分配成为无线通信研究的热点, 资源分配包括子信道分配和功率分配。OFDMA 系统可以充分地利用时间资源, 频率资源和多用户分集。相对于单小区系统, 多小区 OFDMA 系统的资源分配更具有实际意

义，因此显得尤为重要。

在单小区中，对于单用户或多用户的固定的资源分配，在假定信道状况是静态的或慢衰落的且收发端已知的情形下，结合自适应调制的注水算法是最优的^[1]，但注水算法在多小区环境中因为相邻小区之间的共信道干扰而不能保持其最优性。

近年来，许多方法^{[2]-[5]}致力于研究多小区OFDMA系统下行链路的资源分配。文献[2]和[3]采用集中分配的方法或者相邻小区协作来解决这个问题，但这种方法需要额外的信令开销和小区规划。在分布式拓扑结构的共信道干扰管理中，分布式操作是个较好的方法。分布式操作基站不需要来自其它小区的信息。文献[4]和[5]运用非合作博弈去执行分布式资源分配，同时证明了这个博弈可以收敛于唯一的均衡解。但是我们可以注意到应用分布式的非合作博弈不一定会收敛于唯一的纳什均衡解，且纳什均衡解也不一定是系统优化的最优解。

本篇论文提出应用粒子群算法(Particle Swarm Optimization, PSO)优化OFDMA系统的多小区的功率分配问题。粒子群优化算法是由 Kennedy和Eberhart发明的^[6]一种进化算法，利用简单均衡的机制去实现全局和局部搜索。近年来粒子群算法已经成为一个研究热点，是一种应用于非线性优化问题的很好的优化工具。由于粒子群算法并行操作的特点，很适用于分布式功率分配问题的优化。但是目前将PSO应用于功率分配方面尤其应用于OFDMA系统中的工作很少。在文献[7]中，PSO应用于CDMA系统干扰抵消的功率分配的优化。[8]通过有限制条件的PSO解决了无线传感网络的数据融合的最优功率调度问题。文献[9]PSO应用于OFDMA系统的资源分配，但它只解决了子信道分配到用户的优化问题。本篇文章利用粒子群算法优化多小区OFDMA系统的分布式功率分配问题，且假定子信道已经分配完毕。目标是在有基站功率约束的条件下最大化系统的频谱效率。

论文结构安排如下：在第2节，建立系统模型并用公式表述多小区 OFDMA 系统下行链路的优化问题。在第3节，介绍 PSO 技术及基于 PSO 的功率分配问题。第4节展示了仿真结果，第5节分给出结论。

2 系统建模和问题陈述

在这一部分，首先建立系统模型，然后用公式表达功率分配的优化问题。

2.1 系统建模

考虑 N 个小区 OFDMA 蜂窝系统的下行链路。为了提高频谱效率和抑制来自相邻小区的共信道干扰，设每个小区分成三个 120° 的扇区，来自同一小区的三个扇区是由这个小区的基站统一控制。假设所有小区占用共同的频带且总带宽 B 划分为 M 个子信道。在每个扇区内， K 个用户均匀分布，每个子信道只分给一个用户，且每个用户只对应到一个服务基站。为了集中研究功率分配问题，假设子信道已经分配到用户，每个基站传输功率的最大值是 P_{\max} 。假定信道状态信息可以无错误无延迟地反馈给基站。

设 S_n 表示小区 n 的扇区的集合，系统的总扇区数目 $S_T \square S_1 \cup \dots \cup S_N$ 。由系统建模每个小区划分为三个扇区，则 $S_n = S = 3$ ，其中 $n=1,2,\dots,N$ 。取扇区 j ， $j \in S_n$ ， $k(m,j)$ 表示位于扇区 j 通过子信道 m 连接到基站的用户。定义 $G_{i,k(m,j)}^m$ 表示在用户 $k(m,j)$ 和扇区 i 之间在子信

道 m 的信道增益。 p_j^m 表示由基站 n 通过子信道 m 分配到扇区 j 的功率值。 \mathbf{P}_j 表示由基站 n 分配到扇区 j 的 M 维的功率向量， $\mathbf{P}_{(n)} = [\mathbf{P}_1, \dots, \mathbf{P}_j, \dots, \mathbf{P}_{S_n}]^T$ 表示第 n 个小区的功率分配矩阵，是 $S \times M$ 维的。 $\mathbf{P}' = [\mathbf{P}_{(1)}, \dots, \mathbf{P}_{(n)}, \dots, \mathbf{P}_{(N)}]$ 表示这个系统的功率向量。

用户 $k(m, j)$ 的信干噪比(signal-to-noise-and-interference, SINR) $\gamma_{j,k(m,j)}^m(\mathbf{P}')$ ，假设是通过子信道 m 连接到基站 n (取 $j \in S_n$)，可以表示为

$$\gamma_{j,k(m,j)}^m(\mathbf{P}') = \frac{p_j^m G_{j,k(m,j)}^m}{\sum_{\substack{i \in S_n \\ i \neq j}} p_i^m G_{i,k(m,j)}^m + \sum_{\substack{i \in S_r \\ i \notin S_n}} p_i^m G_{i,k(m,j)}^m + \sigma^2} \tag{1}$$

在上式中， $p_j^m G_{j,k(m,j)}^m$ 是有用数据， $\sum_{i \neq j} p_i^m G_{i,k(m,j)}^m$ 是小区内干扰， $\sum_{i \notin S_n} p_i^m G_{i,k(m,j)}^m$ 表示小区间干扰。

用户 k 可达到的数据速率表达为

$$R_k^j(\mathbf{P}') = \sum_{m=1}^M \frac{B}{M} \log_2 \left(1 + \frac{\gamma_{j,k}^m(\mathbf{P}')}{\Gamma} \right) \tag{2}$$

Γ 表示调制间隔， $\Gamma = -\frac{\ln(5BER)}{1.5}$ 是比特误码率 BER[10]。

2.2 问题陈述

目标是在基站发射功率约束的条件下通过优化功率分配问题来最大化系统的吞吐量：

$$\begin{aligned} \max_{\mathbf{P}'} \quad & \tau_n(\mathbf{P}') = \sum_{j=1}^{S_n} \sum_{m=1}^M \frac{B}{M} \log_2(1 + \gamma_{j,k(m,j)}^m) \\ \text{subject to.} \quad & \begin{cases} p_j^m \geq 0 & \forall j, m \\ \sum_{m=1}^M p_j^m \leq P_{\max}, & j \in S_n, \forall S_n \end{cases} \end{aligned} \tag{3}$$

注意到这是一个 NP-hard 问题，这个工作的目标就是优化问题的解。

3 PSO 用于功率分配

首先简要介绍粒子群优化算法，然后详细地描述其在分布式功率分配问题中的应用。

A. 粒子群优化算法

初始化粒子群，定义优化函数为适应函数，那么每个粒子的适应函数的值可以通过将粒子代入适应函数求得。每个粒子在自身更新过程中取得的最优值为个体极值 **pbest**，所有粒子更新过程中取得的最优值为全局极值 **gbest**。每个粒子在不断地检测自己位置的同时跟踪 **pbest** 和 **gbest** 的变化来更新自己的位置。迭代过程中，如果某个粒子的当前适应值比该粒子的当前个体极值所对应的适应值更优，则将个体极值 **pbest** 更新为该粒子的当前位置；否则，其个体

极值保持不变。且在每次粒子更新过程中，每个粒子都将其个体极值与全局极值进行比较，如果前者适应值优于后者，则将全局极值 \mathbf{gbest} 更新为该粒子的个体极值；否则，全局极值保持不变。该更新迭代过程持续进行，直至达到最大迭代次数或粒子充分接近最优值。

B. PSO 用于功率分配

系统内各小区地位平等，每个小区的基站分别根据本小区用户反馈的信道状态信息进行迭代计算，并行地执行粒子群算法寻找本小区最优的功率分配矩阵；分配依据是本小区内不同功率分配矩阵的信道容量差异，并以此判决标准为每个小区选择最优的功率分配矩阵，实现系统信道容量的最大化；因系统内各个小区实现分布式的功率分配，各基站之间不需要进行直接的信息交互，从而在较低的开销下，提高了系统频谱利用率。

算法采用了分布式的并行处理，在此对一个小区的 PSO 用于功率分配优化的过程进行详细的阐述。

1) 定义和初始化

(a) 定义吞吐量函数 $\tau(\mathbf{P})$ 为适应函数，迭代的最大次数设为 K 。在本文功率分配问题中，算法中的粒子表示为本小区的功率分配矩阵，是 $S \times M$ 维的；粒子的更新速度表示功率的更新步长矩阵，是 $S \times M$ 维的。粒子群大小设为 L 。

当系统启动后，基站初始化设置本小区的 L 个 $S \times M$ 维的功率分配矩阵 $(\mathbf{P}^{(1)}, \dots, \mathbf{P}^{(l)}, \dots, \mathbf{P}^{(L)})$ 和 L 个 $S \times M$ 维的更新步长矩阵 $(\mathbf{V}^{(1)}, \dots, \mathbf{V}^{(l)}, \dots, \mathbf{V}^{(L)})$ ，其中 $\mathbf{P}^{(l)}$ 表示第 l 个 $S \times M$ 维的功率分配矩阵， $\mathbf{V}^{(l)}$ 表示第 l 个 $S \times M$ 维的更新步长矩阵， l 为不大于 L 的正整数。

初始化设置基站的每个功率分配矩阵 $\mathbf{P}^{(l)}$ ($l=1,2,\dots,L$)，使之服从 $(0, \frac{2 \times P_{\max}}{S \times M})$ 的均匀分布，以计算同频干扰和获取较好结果；功率分配矩阵中的各个元素均为非负数值，且所有元素之和不大于 P_{\max} 。初始化设置每个更新步长矩阵 $\mathbf{V}^{(l)}$ ($l=1,2,\dots,L$) 为零。

(b) 对于第 l 个功率分配矩阵 $\mathbf{P}^{(l)}$ ，计算 S 个扇区的信道容量之和 τ_l ，并以 τ_l 作为 $\mathbf{P}^{(l)}$ 的评判依据；将第 l 个功率分配矩阵在迭代过程中获得的当前的最高信道容量之和 $\tilde{\tau}_q^l$ 的矩阵定义为其当前最优矩阵 $\bar{\mathbf{P}}_q^l$ ；依此类推， l 取 $1,2,\dots$ 到 L ，逐个计算 L 个功率分配矩阵的当前最优矩阵，并用 $\bar{\mathbf{P}}_q = \{\bar{\mathbf{P}}_q^1, \bar{\mathbf{P}}_q^2, \dots, \bar{\mathbf{P}}_q^L\}$ 表示之，该初始化设置为全零矩阵；

将迭代结束时 L 个当前最优矩阵中的对应最高信道容量之和的矩阵定义为全局最优矩阵，用 $\bar{\mathbf{P}}_g$ 表示。粒子群算法迭代的目标即选取全局最优矩阵作为本小区的功率分配矩阵。初始化设置 $\bar{\mathbf{P}}_g$ 矩阵为零；全局最优矩阵 $\bar{\mathbf{P}}_g$ 对应的最高信道容量之和记为 $\tilde{\tau}_g$ 。

2) 迭代更新过程

(a) 基站根据用户终端的反馈信息收集各个扇区在每个子信道上的信道增益矩阵 \mathbf{G} 与干扰矩阵 \mathbf{T} ，矩阵 \mathbf{G} 和 \mathbf{T} 均为 $S \times M$ 维矩阵。

对于第 l 个功率分配矩阵 $\mathbf{P}^{(l)}$ ，根据收集的信道增益信息和干扰信息以及功率分配矩阵 $\mathbf{P}^{(l)}$ ，计算该小区的信道容量之和 τ_l' ，如果新的信道容量之和 τ_l' 大于当前最高信道容量之和 $\tilde{\tau}_q^l$ ，则更新 $\tilde{\tau}_q^l = \tau_l'$ ，并且更新 $\mathbf{P}^{(l)}$ 为第 l 个功率分配矩阵的当前最优矩阵 $\bar{\mathbf{P}}_q^l = \mathbf{P}^{(l)}$ ；反之， $\bar{\mathbf{P}}_q^l$ 和对应的 $\tilde{\tau}_q^l$ 均保持不变。

根据当前时刻更新的 L 个功率分配矩阵的当前最优矩阵 $\bar{\mathbf{P}}_q = \{\bar{\mathbf{P}}_q^1, \bar{\mathbf{P}}_q^2, \dots, \bar{\mathbf{P}}_q^L\}$ ，其对应的更

新的小区信道容量之和为 $\tilde{\tau}_q = \{\tilde{\tau}_q^1, \tilde{\tau}_q^2, \dots, \tilde{\tau}_q^L\}$ ，设 τ_g' 为 $\tilde{\tau}_q = \{\tilde{\tau}_q^1, \tilde{\tau}_q^2, \dots, \tilde{\tau}_q^L\}$ 中的最大值，且设其对应的是第 k 个功率分配矩阵， k 的取值范围是 $[1, L]$ ；再判断该新的全局最高信道容量之和 τ_g' 是否大于当前全局最优矩阵 $\bar{\mathbf{P}}_g$ 对应的最高信道容量之和 $\tilde{\tau}_g$ ，如果 τ_g' 大于 $\tilde{\tau}_g$ ，则更新 $\tilde{\tau}_g = \tau_g'$ ，并将 $\mathbf{P}^{(k)}$ 更新为全局最优矩阵，即 $\bar{\mathbf{P}}_g = \mathbf{P}^{(k)}$ ；反之， $\bar{\mathbf{P}}_g$ 和对应的 $\tilde{\tau}_g$ 都保持不变。

(b) 更新该小区所有的更新步长矩阵 \mathbf{V} 和功率分配矩阵 \mathbf{P} ：将更新步长矩阵更新为 $\mathbf{V}^{(l)} = w \times \mathbf{V}^{(l)} + c_1 \times r_1 \times (\bar{\mathbf{P}}_q^l - \mathbf{P}^{(l)}) + c_2 \times r_2 \times (\bar{\mathbf{P}}_g - \mathbf{P}^{(l)})$ ，式中， w 为可调常数，用于表示更新步长矩阵受前一时刻的影响，可调参数 r_1, r_2 是两个分别依照设定规律生成的随机数，权重因子 c_1 和 c_2 代表速度分别受当前最优矩阵和全局最优矩阵的影响程度；两个乘积 $c_1 \times r_1$ ， $c_2 \times r_2$ 分别代表更新步长矩阵受当前最优矩阵和全局最优矩阵的影响程度，上述参数的设置会在很大程度上影响到运用粒子群算法迭代寻找最优功率分配矩阵的性能；

再将功率分配矩阵更新为 $\mathbf{P}^{(l)} = \mathbf{P}^{(l)} + \mathbf{V}^{(l)}$ ，因矩阵 $\mathbf{V}^{(l)}$ 为非负矩阵，则更新时能够满足 $\mathbf{P}^{(l)}$ 的非负性能；如果 $\mathbf{P}^{(l)}$ 中所有元素的和 P_{sum} 大于基站的最大发射功率 P_{max} ，则 $\mathbf{P}^{(l)}$ 中所有的元素都缩小 $\frac{P_{max}}{P_{sum}}$ 倍，即 $\mathbf{P}^{(l)} = \frac{P_{max}}{P_{sum}} \times \mathbf{P}^{(l)}$ ；

(c) 继续进行迭代操作。判断如果达到迭代的最大次数，结束该操作流程，进入步骤 (3)；判断如果没有达到迭代终止条件，重复操作步骤 (2)，继续更新当前最优功率分配矩阵和全局最优功率分配矩阵。尽可能地实现该小区系统频谱效率或信道容量之和的最大化。

4 输出该小区的最优功率分配方案 $\bar{\mathbf{P}}_g$

PSO 的一个显著的特点是简单性，它只需要两个公式，和几个参数的调整。

参数的设置极大得影响 PSO 的性能。设置影响收敛概率和收敛速度，因此参数的选择显得尤为重要。像其它的启发式算法一样，PSO 中没有明确的方法去设置参数，且没有保证可以收敛到最优值。

本文中，设置粒子群规模 $L = 30$ ，这个群体大小适于大多数的电子类问题 [11]。研究者发现越大的惯性指数利于跳出局部最优解，而越小的惯性指数利于算法的收敛。因此，一个方法就是将 w 随迭代次数线性减小，如由 1.4 到 0 递减，或由 0.95 到 0.2 递减。Eberhart 和 Shi 建议 w 可以由 0.9 线性到 0.4。 c_1 和 c_2 设置为常数，但不一定为 2。

算法的时间复杂度以迭代次数来测量。在 PSO 的优化过程中，每个基站独立运作且算法是并行迭代的过程，可以证明在每次迭代过程中，对于每个基站，其 PSO 优化的复杂度是 $O(s \times M)$ ， s 表示每个小区扇区的数目， M 表示迭代次数。最大迭代次数 K 与粒子群的大小有关。

5 仿真结果

通过建立模型进行仿真，来评价所提方法的性能。

A. 信道模型和多小区结构

在 OFDMA 蜂窝系统 7 小区的环境中，每个小区划分为 3 个 120° 的协作的山区。路径损耗模型是车载环境。 $L=128.1+37.6\log_{10}(d)$ ，载频在 2MHz 附近。大尺度衰落为零均值，方差为 8dB 的对数正态分布。小尺度为六径衰落模型，功率依次是(0.0,-1.0,-9.0,-10.0,-15.0,-20, 0) dB,时延为(0, 310, 710, 1090, 1730, 2510) ns。多普勒频移遵循最大多普勒频移为 100Hz 的 Clarke 模型。总带宽 B 是 10MHz，子信道（资源块）的总数目是 $M=50$ 。

设置参数如下：小区半径，0.5Km;每扇区的用户数为 50，用户在基站附近均匀分布，最小距离为 0.035km；所有扇区的最大传输功率为 $P_{\max}=10\text{W}$ ；噪声功率谱密度，-174dBm/Hz; 目标 BER， 10^{-4} 。基站方向天线模式 $A(\theta)=-\min\left[12(\theta/\theta_{\text{3dB}})^2,A_m\right]$ ，其中 $\theta_{\text{3dB}}=70^\circ$ ， $A_m=20\text{dB}$ 。仿真时间为 100 TTI (Time Transmission Interval)，时间传输间隔 TTI 为 0.5ms。

设置粒子群大小 $L=30$ ，粒子群初始化可以遵循在范围(0, $2P_{\max}/M$)内的均匀分布，均值为 P_{\max}/M 。 $c_1=c_2=2$ ， γ_1 服从 (0, 1) 的均匀分布， γ_2 从 0.8 线性增加到 1.0，其线性增加的步长可调；设 w 从 1.2 线性递减至 α ，其中 $0.8<\alpha<1$ ，线性递减的步长同样可调。

B. 仿真图示

我们把本文提出的 PSO 的方法同另外两种采用相同系统模型的方法比较。对比的第一种方法是平均分配功率方法。第二种是在每个扇区执行功率注水方法。下图展示了在 7 小区系统中采用三种方法的仿真结果。

图 1 表示所提 PSO 的优化方法比另外两种方法在系统的频谱效率方面性能提高很大。平均来说，PSO 比平均分配功率方法提高了 54%；PSO 比注水方法提高了 30%。

图 2 展示了采用 PSO 优化功率分配问题的适应函数的收敛曲线。粒子群大小为 30，接近收敛的迭代次数为 200。图示显示在第 100 次的迭代数据达到了最后结果的 98.5%，如果需要可以减小迭代次数。

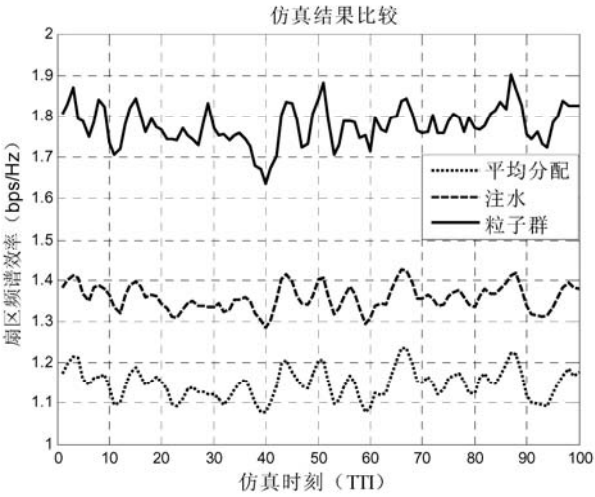


图 1 7 小区的频谱效率仿真结果

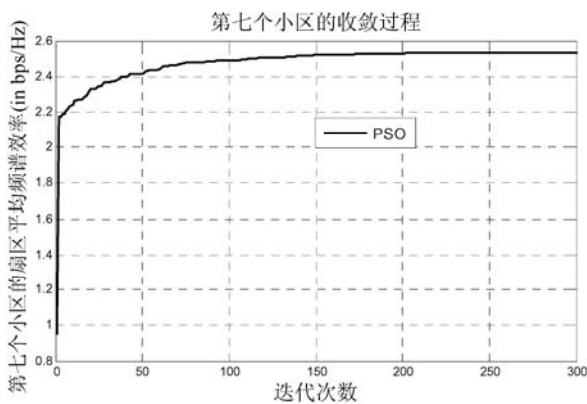


图 2 PSO 方法的收敛曲线

6 结论

本文提出在多小区 OFDMA 蜂窝系统中采用粒子群优化算法的分布式功率分配的方法。目标是在有基站传输功率约束的条件下最大化系统的频谱效率。与注水方法相比较,提出的方法实现了小区内干扰协调,极大地提高了系统的频谱效率。由仿真结果显示,在 7 个小区的环境下提出的分布式的方法在系统频谱效率方面分别比平均分配功率方法和注水方法提高了 54%和 30%。

参 考 文 献

- [1] W. Rhee and J. M. Cioffi, "Increase in capacity of Multiuser OFDM system using dynamic sub-channel allocation[C]," IEEE VTC'2000, pp. 1085 - 1089 Tokyo, Japan, May 2000
- [2] G. Li and H. Liu, "Downlink dynamic resource allocation for multi-cell OFDMA system[C]," in Proc. IEEE VTC 2003, Orlando, Oct. 2003.
- [3] J. Li, H. Kim, Y. Lee and Y. Kim, "A novel broadband wireless OFDMA scheme for downlink in cellular communications[C]," in Proc. IEEE WCNC 2003, New Orleans, Mar. 2003.
- [4] Zhu Han, Zhu Ji, and K. J. Ray Liu, "Non-Cooperative Resource Competition Game by Virtual Referee in Multi-Cell OFDMA Networks[J]," IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS, VOL. 25, NO. 6, AUGUST 2007
- [5] Hojoong Kwon and Byeong Gi Lee, "Distributed Resource Allocation through Non-cooperative Game Approach in Multi-cell OFDMA Systems[C]," in proc. IEEE ICC 2006 .
- [6] J. Kennedy and R. Eberhart, "Particle swarm optimization [C]," Proc. IEEE Int. Conf. Neural Networks, 1995, Vol. IV, pp. 1942-1948.
- [7] K. Zielinski, P. Weitkemper, R. Laur, K.-D. Kammeyer: "Optimization of Power Allocation for Interference Cancellation with Particle Swarm Optimization[J]", IEEE Transactions on Evolutionary Computation, 13(1):128-150, 2009

- [8] Wimalajeewa, T.; Jayaweera, S.K. "Optimal Power Scheduling for Correlated Data Fusion in Wireless Sensor Networks via Constrained PSO[J], " IEEE Transactions on Wireless Communications, vol.7, pp.3608-3618,Sep.2008
- [9] Shahin.Gheitanchi, Falan Ali,Elias Stipidis, "Particle Swarm Optimizaion for Resource Allocation in OFDMA[C]", in Proc. IEEE. Int. Conf. Digital Signal Proceeding. pp. 383-386.Jul.2007
- [10] A.J.Goldsmith and S.-G. Chua"Variable-rate variable-power MQAM for fading channels[J]," IEEE Transaction on Communications vol. 45, pp. 1218–1230, Oct. 1997.
- [11] A. Carlisle and G. Doizier, "An off-the-shelf PSO[C]", Proc. 2001 Workshop on Particle Swarm Optimization, 2001, Indianapolis, IN

作者简介:

孙文文, 女(1985-), 硕士研究生, 无线通信系统中的资源分配

贺志强, 男(1975-), 副教授, 信息理论与移动通信

牛凯, 男(1976-), 副教授, 信息理论与移动通信

田宝玉, 男(1946-), 教授, 数字移动通信理论与技术

一种适合于低压PLC的组合译码HARQ算法

朱子善 林 东

(福州大学 物理与信息工程学院; 福建 福州 350002)

摘 要: 针对低压 PLC 信道干扰大且具有时变性的特点, 提出一种采用组合译码的 HARQ 方案。通过把 ARQ 过程中多次重传的历史数据进行组合译码获得可信度较高的新码字从而可能提前结束重传过程。计算机仿真结果表明采用该算法的 HARQ 在有效减少重传次数及改善误码性能方面明显优于一般的 HARQ。

关键词: 通信技术; HARQ; 组合译码; 电力线通信

A Combined Decoding Algorithm for Low Power PLC Base on Hybrid ARQ

Zhu Zi shan Lin Dong

(College of Physics and Information Engineering, Fuzhou University, Fuzhou Fujian 350002, China)

Abstract: An algorithm of combined decoding base on hybrid ARQ was presented considering the strong interference and time variability of low power PLC channel. The repeated packets were combined to obtain a new packet that can be decoded successfully and the retransmission ended earlier. The computer simulation indicates that the HARQ with the new combined decoding algorithm is obviously superior to the normal HARQ in terms of the number of retransmissions and bit error rate.

Keywords: communication technology; HARQ; combined decoding; PLC

0 引言

电力线载波通信(Power Line Communication, 简称 PLC)利用现有的电力线, 通过载波方式传输模拟或数字信号。PLC 用于信号传输的介质已经存在, 具有不用重新布线、覆盖范围广、连接方便、安装便捷、功能灵活、扩展容易等显著特点, 有着其他媒介不能相比的优势。随着各种抗干扰通信技术的出现和发展, PLC 应用越来越广。目前该技术开始应用于家居自动化、远程抄表、宽带上网等领域。

但是电力线本身是用来传输电能, 作为通信信道其传输特性很差, 突出的问题是工作环境较恶劣、信号衰减强、干扰大且具时变性^[1]。信号衰减大主要是负载对信号传输影响很大, 如容性负载对载波信号几乎就是短路; 时变性主要是由于电网上负载的不断接入、切出等随机事件的发生; 干扰大且复杂, 可以分为周期性和非周期性的, 其主要来源有谐波噪声(频率为交流电频率的倍数)、大量用电设备发出的表现为低频调幅性的干扰和一些随机的干扰如电

子开关操作等等。对数据通信系统有很重要影响的是脉冲干扰，它能产生突发性干扰引起瞬间的高误码率，这些干扰持续时间短，强度大小不等，随机出现，但其能量高、宽频谱的特性，使所传送数据的若干位甚至整个数据传输过程发生错误。

针对电力线特有的信道特性，必须有一个较好的技术来提高信号传输的可靠性，减小误码率，增加数据的传输距离。虽然在电力线通信中可以采用 OFDM、扩频等技术来增强抗干扰问题^[2]，但是实际系统中还必须利用一定的差错控制技术来提高系统的抗干扰能力，增强传输的可靠性。另一方面低压 PLC 的通信设备必须满足小型，低成本的要求，所采用的差错控制方式应具有算法简单，实用可靠的特点。本文提出一种适合低压 PLC 的采用组合译码的 HARQ 算法，通过把 ARQ 过程中多次重传的历史数据进行组合译码获得可信度较高的新码字从而可能提前结束重传过程，有效减少重传次数，改善误码性能。

1 混合ARQ 简介

最常采用的差错控制方式有前向纠错(FEC)和自动检错重发(ARQ)。前向纠错方式采用的编码方式决定其编码效率和纠错能力。在传输信道误码率较高时，必须使用较低编码效率、纠错能力强的码组，才能够提供足够高的系统可靠性，但是这使得编译码器的复杂度过高，系统吞吐量偏低。由于低压 PLC 信道特性不稳定，通常仅依赖前向纠错码还难以保证足够的系统可靠性，因而不宜单独采用。ARQ 方式传输次数随信道条件的变化而自动调整，适宜于在信道特性不稳定的通信系统中采用。但低压电力线载波信道传输特性特别不稳定，信道特性衰落持续时间较长，即使采用 ARQ 方式，多次重传也仍有可能得不到正确码字，这也是形成远程抄表盲区现象的主要原因之一。

混合 ARQ(Hybrid ARQ, HARQ)结合 FEC 和 ARQ 两种差错控制方案的优点。HARQ 可以看成是在传统的 ARQ 系统中引入一个 FEC 子系统，在纠错码纠错能力范围内的错误利用前向纠错码纠正，从而降低 ARQ 系统的重传次数；而对于前向纠错码无法纠正的错误，检错码通常可以检测出错误，并要求 ARQ 系统重传该分组。在信道条件较好时有合理的传输效率，又能在信道条件变差时保证数据的可靠性和通过率，因此 HARQ 可进一步提高系统传输可靠性和吞吐量。

但是低压 PLC 的信道中存在强脉冲干扰，信道的比特差错率高，当误码率很高时，HARQ 重发过于频繁而使效率大大降低，传输效果仍不理想。针对这类系统，可利用码字组合译码技术对 HARQ 进行改进，提高系统的译码成功率。

2 组合译码算法

组合译码的概念最初是由 David Chase 提出的^[3]，码字组合译码实际上是一种时间上的分集接收技术。它利用 ARQ 系统中对同一数据进行多次重复传送这一特点，将各次传输的结果进行最佳加权组合后再进行译码。在一定条件下有可能从收到的多个有错码字中译出正确码字，从而结束重传过程^[4]。

其基本原理是对码字重复发送多次得到一个编码效率低于原码字的新码字。对该新码字进行最大似然译码，可得到高于原码字的纠错能力。

设发端发送码字的码长为 N ，其中信息比特有 k 位，则可用码字数为 2^k 。

发端发送的码字记为 $X = x_1 + x_2 + \cdots + x_N$ ，

收端收到的码字记为 $Y = y_1 + y_2 + \cdots + y_N$ 。

对码字 X 重复发送 M 次，收端收到的 M 个码字依次记为 $Y^1, Y^2, \cdots Y^M$ 。

在收端将这 M 个码字组合成一个码字进行最大似然译码，就是在所有 2^k 个可用码字 $X_i = x_{i1} + x_{i2} + \cdots + x_{iN}$ ($i = 1, 2, 3, \cdots 2^k$) 中选择出接收的 M 个码字 Y 与发送码字 X 之间条件概率最大时 i 所对应的码字作为译码结果，条件概率按下式计算：

$$p[Y | X_i] = \prod_{m=1}^M (1 - p_m)^{N - d_{mi}} p_m^{d_{mi}} \quad (1)$$

其中 d_{mi} 是第 m 次传输 Y^m 与 X_i 不一致比特的位数， p_m 是传输 Y^m 时的比特差错率，由第 m 次传输时信道的情况决定。例如可通过对预先已知的同步字的错码个数进行统计来计算 p_m ，也可通过解调器输出的信噪比指标来估计 p_m 。一般来讲，低压 PLC 信道是不稳定的， p_m 是时间的函数。但当一个码字的传输时间很短时，可以认为 p_m 为一个常数。

对式(1)两边取自然对数得：

$$\max_i \{\ln p[Y | X_i]\} = \max_i \left\{ \sum_{m=1}^M N \ln(1 - p_m) - d_{mi} \ln\left(\frac{1 - p_m}{p_m}\right) \right\} \quad (2)$$

由于 $N \ln(1 - p_m)$ 是与 i 无关的函数，式(2)可改写为：

$$\max_i \{\ln p[Y | X_i]\} \propto \min_i \left\{ \sum_{m=1}^M d_{mi} \ln\left(\frac{1 - p_m}{p_m}\right) \right\} \quad (3)$$

$\ln\left(\frac{1 - p_m}{p_m}\right)$ 为可靠性因子，表征了第 m 次传输时信道的可信度。 p_m 愈大， $\ln\left(\frac{1 - p_m}{p_m}\right)$ 愈小。

这实际上是对各次传送的码序列进行加权运算。而 p_m 要由信道的当时特性来决定。例如可通过对预先已知的同步字的错码个数进行统计来计算 p_m ，也可通过解调器输出的信噪比指标来估计 p_m 。尤其是在对同一码字的多次重发期间信道噪声水平变化不大时，则可将 $\ln\left(\frac{1 - p_m}{p_m}\right)$ 项略去。

对于码长为 N 的二进制码，有

$$d_{mi} = \sum_{j=1}^N x_{ij} \oplus y_j^m \quad (4)$$

代入式(3)有

$$\max_i \{\ln p[Y | X_i]\} \propto \min_i \sum_{m=1}^M \ln\left(\frac{1 - p_m}{p_m}\right) \sum_{j=1}^N x_{ij} \oplus y_j^m \quad (5)$$

这实际上是将收到的 M 个码字逐一与所有可用码字进行比较，再将每次比较的结果进行加权求和，和值最小的 i 所对应的可用码字与收到的 M 个码字的综合距离最小，所以将该码字作为译码结果，在所有可用码字出现的概率相等的条件下，是最佳译码。

由于原码字的编码效率为 k/N ，重复 M 次得到的组合码的编码效率为 $k/(NM)$ ，所以只需增加重复次数就可使信息速率低于香农信息论所定义的信道容量从而使得无差错传输成为可能。但是，由式(5)可知，当原码字的码长，尤其是信息比特数较多时，译码的计算量很大，且随着重发次数的增加而增加。为此可考虑分步译码算法。第 1 步按比特进行最大似然译码，

译码公式为

$$Y_{cj} = \min \sum_{m=1}^M (y_j^m \oplus x_{ij}) \ln \left(\frac{1-p_m}{p_m} \right) \tag{6}$$

对于二进制码， $X_{ij} \in (0, 1)$ ，式(6)可简化为

$$Y_{cj} = \min \left\{ \sum_{m=1}^M (y_j^m \oplus 1) \ln \left(\frac{1-p_m}{p_m} \right), \sum_{m=1}^M (y_j^m \oplus 0) \ln \left(\frac{1-p_m}{p_m} \right) \right\} \tag{7}$$

这实际上是比较在 M 次传输中，第 j 位为 1 和为 0 的次数。若为 1 的次数多，则将第 j 位译为 1；反之，则译为 0，运算量大大减少。且运算是一个叠加累计过程，不必存储收到的每一个码字，可节省接收端的存储空间。

当 $X_{ij}=1$ 和 $X_{ij}=0$ 的概率相等时，式(7)的译码是最佳的。将对 N 比特进行译码所得结果依次排列，即可得到一个新码字，记为 Y_c 。 Y_c 是按比特译码得到的，它比任一个 $Y^m \in (Y^1, Y^2, \dots, Y^M)$ 更接近于原码字，但并不能保证它是原码字。接着对 Y_c 进行检错译码，若检不出错误，则认为组合译码成功。若检出错误，则认为译码失败，要求重发。

首先考虑最简单的情况。对同一码字的重传期间信道特性变化不大时，可将公式(7)中的可靠性因子省略。则按比特译码的公式简化为

$$Y_{cj} = \min \left\{ \sum_{m=1}^M (y_j^m \oplus 1), \sum_{m=1}^M (y_j^m \oplus 0) \right\} \tag{8}$$

利用式(8)译码，只有在 M 次传输中有半数以上在同一比特位置出错时，该比特才会被错译。显然，即使每次重发都发生错码，将所有有错码字组合后进行译码，仍可能得到正确结果。

3 采用组合译码的 HARQ 方案

在一般的 HARQ 系统中，FEC 在纠错范围内进行纠错，否则就要请求重传。采用组合译码的 HARQ 系统，在 ARQ 过程中当错误比特超过 FEC 纠错范围时可先尝试进行组合译码，当组合译码仍有错时再请求重传。采用组合译码算法的 HARQ 系统发送端与普通的 HARQ 发端基本相同，算法处理在接收端进行。

一个采用组合译码的 HARQ 系统使用的信息分组可由如下所示的三部分组成：

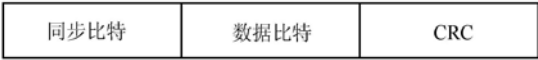


图 1 信息分组结构

同步比特用于检测信息分组的存在和指示接收数据的起始位置，同样可以用于估计此时信道的可信度，获得该分组的可靠性因子。

数据比特为实际要传输的信息比特，也可包含一定的网络传输控制信息。CRC 比特用于循环冗余检验，判断该信息分组是否译码正确。数据比特和 CRC 比特都必须经过纠错编码。当编码解码必须以软件方式在低成本微控制器中实时地执行时，基于汉明码的分组编码通常是首要的选择^[5]。在该方案中采用(7, 4)汉明编码，能在一个码字内纠正一个错误或检测两个错误。

记信息分组 S 第 m 次($m=1, 2, \dots, M$)经 CRC 校验和汉明码编码后由 PLC 信道传输，在接收端接收到的信息分组为 $R(m)$ ， $R(m)$ 的汉明码译码信息输出为 $D(m)=\text{Hamming_D}(R(m))$ ，CRC

校验给出接收分组有无差错指示 IE(IE=0 无错, 否则有错), 记为 $IE=CRC_D(D(m))$, 由 $R(1) \dots R(m)$ 按式(7)组合译码输出记为 R_C , 接收端最终输出信息分组为 \hat{S} , 则接收端采用组合译码算法的 HARQ 过程可由图 2 描述。

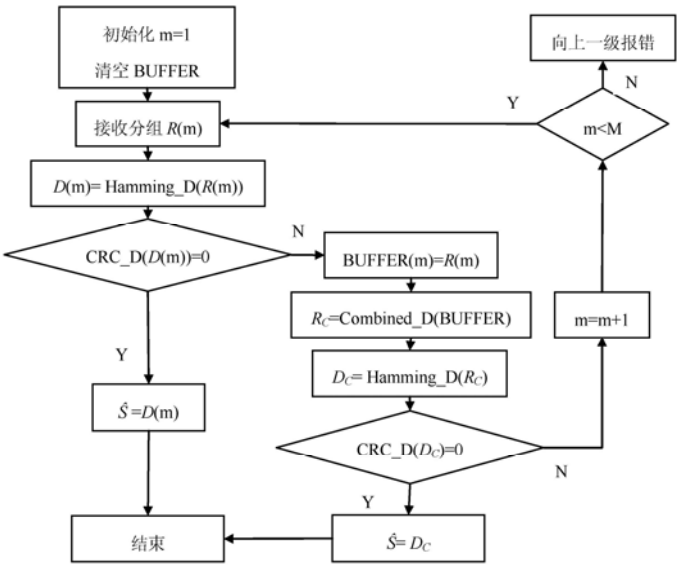


图 2 接收端采用组合译码的 HARQ 过程

4 仿真结果及分析

利用 MATLAB 分别对本文提出的 HARQ 过程和一般 HARQ 过程在加性高斯白噪声环境下进行计算机仿真, 采用 (7, 4) 汉明码, BPSK 调制, 最大允许传输次数 (含首传) 分别为 3、5、7 次, 考虑到在实际 PLC 系统一般为发射功率受限的系统, 对信号和噪声功率比 (SNR) 不作 E_b/n_0 的归一化, 而采用信道实际传输过程中的 SNR 作为相互比较的基准。图 3 给出了这两种 HARQ 系统的误码率曲线。在低信噪比条件下, 本文 HARQ 方案的抗干扰能力明显优于一般 HARQ 方案, 且允许最大传输次数越大, 在相同的信道 SNR 下可以获得更低的误码率。

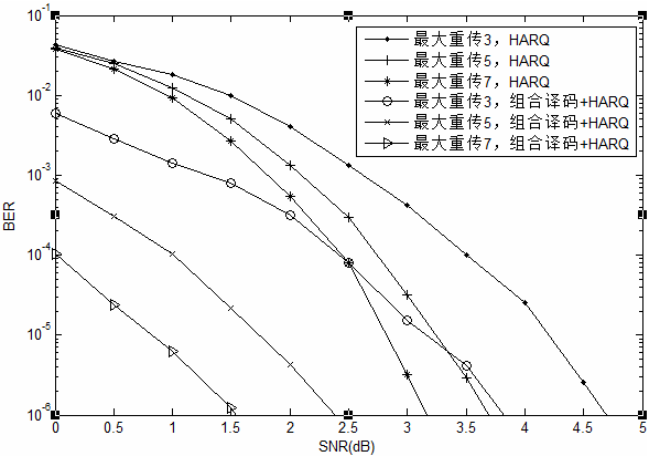


图 3 组合译码 HARQ 与无组合译码 HARQ 系统误码性能比较

图 4 给出了平均传输次数性能比较图。当最大允许传输次数为 3 时，两种 HARQ 系统的平均传输次数基本相同，随着最大允许传输次数的增大，在较低信噪比时采用组合译码算法的 HARQ 平均分组传输次数与一般 HARQ 平均传输次数相比有明显改善，当信道 SNR 小于 2dB，最大重传次数为 7 时平均传输次数减少了 40%~50%，说明采用组合译码算法的 HARQ 系统在信道较差时能更早结束重传过程，与一般 HARQ 相比可获得更好的吞吐率。

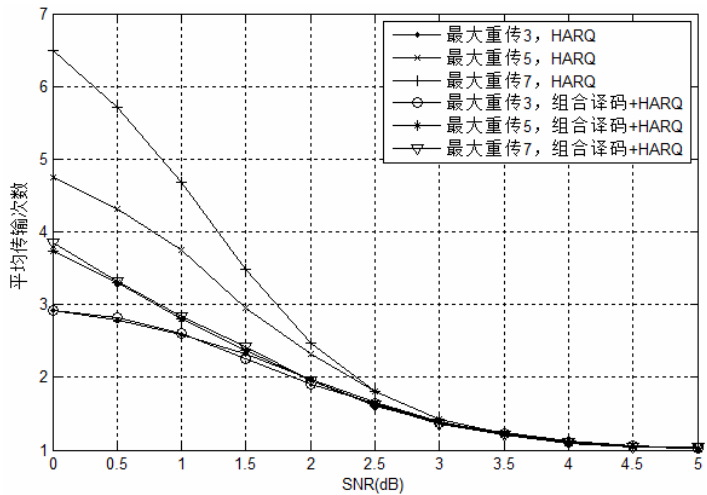


图 4 平均传输次数性能比较

5 结束语

采用组合译码算法的 HARQ 将 ARQ 过程中多次重传的历史数据加以利用，从中组合出一个可信度较高的新码字，既可降低平均重传次数，又提高了抗干扰能力，能很好的改善实际应用中发射功率受限系统在信噪比较低时的性能，特别适用于低压 PLC 信道等干扰强的通信环境。

参 考 文 献

[1] 高峰,董亚波.低压电力线载波通信中信号传输特性分析[J].电力系统自动化,2000,24(7):36-40.
[2] 曹宁,胡弘莽.电网通信技术[M].北京:中国水利水电出版社,2003
[3] CHASE D. Code combining—A maximum_likelihood decoding approach for combining an arbitrary number of noisy packets [J].IEEE Transactions on Communications, 1985,33(5):385-393.
[4] 王振朝.利用组合码进行纠错的检错重传方案[J].遥测遥控,1990,11(2):25-29.
[5] Klaus Dostert 著,栗宁,郑福生译.电力线通信[M].北京:中国电力出版社,2003

一种基于信道可靠性的协作频谱感知算法

岳文静¹ 郑宝玉^{1,2}

(1. 上海交通大学电子工程系, 上海 200240;

2. 南京邮电大学 信号处理与传输研究院, 江苏 南京 210003)

摘要: 由于协作频谱感知能够更准确的确定主用户的存在与否, 因此, 它被认为是认知无线电技术中的一个重要研究分支。在协作频谱感知中, 各认知用户需要通过报告信道发送其本地检测结果 (0 或 1) 到中心接入点 (AP), 然而, 在实际中, 报告信道通常是受衰落影响的, 这将降低频谱感知算法的检测性能。针对这个问题, 提出了一种基于信道可靠性的协作频谱感知算法。该算法通过减小由衰落信道引进的误码概率来提高协作频谱感知的检测性能。其基本思想为: 只有具有较好瞬时信道增益报告信道的认知用户才将其本地二进制判决结果发送到 AP, 否则, 该认知用户不发送其本地检测结果到 AP。最后, 对提出的基于信道可靠性的协作频谱感知算法的检测性能进行理论分析研究, 并对其进行了数值仿真。仿真结果表明: 相比于传统的协作频谱感知, 提出的这种协作频谱感知算法能够大大提高对主用户的检测性能。

关键词: 认知无线电; 协作频谱感知; 能量检测; 动态频谱接入

Channel Reliability Based Cooperative Spectrum Sensing Algorithms for Fading Channels in Cognitive Radio Systems

Yue Wen Jing¹ Zheng Bao Yu^{1,2}

(¹ Department of Electronic Engineering, Shanghai Jiaotong University, Shanghai 200240, China;

² Institute of Signal Processing and Transmission, Nanjing University of Posts and Telecommunications, Nanjing, Jiangsu 210003, China)

Abstract: In cognitive radio systems, cooperative spectrum sensing can detect the presence of the primary user with more accurately. However, since the sensing observations are forwarded to an access point (AP) through fading channels, the sensing performance is severely degraded. To deal with this problem, in this paper, a reliability based cooperative spectrum sensing method is proposed to improve sensing performance by decreasing the reporting errors introduced by the fading channels. More specifically, only the secondary users with better instantaneous SNR for reporting channels are allowed to send their local binary decisions (0 or 1) to the AP while the others will send nothing. By employing methodology proposed, the reporting errors introduced by the fading channel can be decrease and then the sensing performance is improved significantly. For proposed reliability based cooperative spectrum sensing, the sensing performance is derived and the analytical performance results are given. Our analysis and numerical results verify that better performance can be achieved

in proposed method compared to conventional cooperative spectrum sensing with OR-Rule.

Keywords: Cognitive Radio, Cooperative Spectrum Sensing, Energy Detection, Dynamic Spectrum Access

为了处理频谱资源紧缺和频谱低利用率之间的矛盾, 认知无线电 (Cognitive Radio, CR) 技术被认为是解决目前无线授权频谱低利用率的最佳方案^[1]。而频谱感知是认知无线电中非常重要的一个研究课题^[2]。一般而言, 单个认知用户的频谱感知算法容易受到多径和阴影效应的影响, 从而使认知用户不能准确地确定未被占用的频段。而协作频谱感知是解决上述问题的有效方法之一。

协作频谱感知通常分为感知 (Sensing) 和报告 (Reporting) 两个阶段。在感知阶段, 每个 SU 独立的完成本地检测; 在报告阶段, 将所有 SUs 的本地检测结果发送到中心接入点 (Access Point, AP), AP 对收到的结果进行数据融合, 综合作出主用户存在与否的最终判决。通过对文献[3]的分析可知, 协作频谱感知可以大大提高对主用户的检测。然而, 协作频谱感知有个共同的问题就是: 要求 SUs 将其本地检测结果发送到 AP 的链路 (即报告信道) 必须是可靠的。到目前为止, 绝大多数有关协作频谱感知的研究主要是集中在本地的检测结果通过理想的报告信道 (无错传输) 发送到 AP。然而, 在实际中, SUs 与 AP 之间的链路不可能是理想的, 因为它往往受到衰落信道的影响。当考虑由报告信道的衰落特性引进的误码时, 这将明显的降低协作频谱感知的检测性能。针对上述问题本文提出一种基于信道可靠性的协作频谱感知方法。该算法的基本点是选择满足要求的认知用户进行协作感知, 某认知用户是否参与协作取决于该用户到 AP 链路的瞬时信道增益。这是因为: 如果本地检测结果通过具有较高信噪比 (SNR) 的报告信道时, AP 接收到的该结果就越可信。因此, 如果第 i 个次用户到 AP 链路的瞬时信噪比 γ_i 大于预先设置的信噪比门限值 γ_T 时, 我们就让该 SU 发送其本地检测结果到 AP, 否则, 该 SU 不发送本地检测结果到 AP。我们分析了提出的这种协作频谱感知算法的性能, 同时仿真也表明提出的方法能大大提高对主用户的检测性能。

1 基于能量检测器的协作频谱感知

我们考虑的认知网络模型由 N 个认知用户和一个中心接入点 (AP) 构成, 如图 1 所示。对于频谱感知问题可以描述为二元假设检验问题。 H_0 对应着 PU 信号不存在; H_1 对应着 PU 信号存在。

首先考虑每个认知用户的本地检测算法。能量检测器由于低计算复杂度成为目前常用的一种检测算法[4]。对于第 i 个 SU 的能量检测器的输出 E_i 服从分布

$$E_i \sim \begin{cases} \chi_{2TW}^2 & H_0 \\ \chi_{2TW}^2 (2\rho_i) & H_1 \end{cases} \quad (1)$$

其中, χ_{2TW}^2 和 $\chi_{2TW}^2 (2\rho_i)$ 分别表示中心和非中心卡方分布, 每个具有 $2TW$ 个自由度, 同时 $2\rho_i$ 是一个非中心参数。 ρ_i 是第 i 个认知用户的瞬时信噪比。为了简单起见, 假设时间带宽

积 TW 为整数, 用 m 来表示。

对于给定瞬时信噪比 ρ_i 和判决门限值 λ , 在 AWGN 信道下, 第 i 个用户的虚警概率 $P_f^{(i)}$ 和检测概率 $P_d^{(i)}$ 分别可以写为^[5]

$$P_f^{(i)} = \Gamma(m, \lambda/2) / \Gamma(m) \quad \text{和} \quad P_d^{(i)} = Q_m(\sqrt{2\rho_i}, \sqrt{\lambda}) \quad (2)$$

其中, $\Gamma(\cdot)$ 和 $\Gamma(\dots)$ 分别为完全和非完全 gamma 函数, $Q_m(\dots)$ 为广义 Marcum Q 函数。

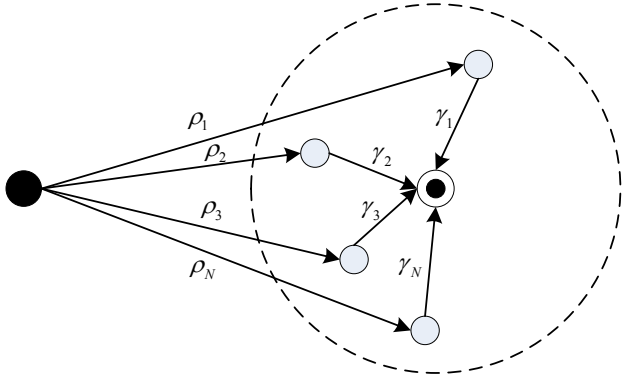


图 1 协作频谱感知模型

在瑞利信道下, $P_f^{(i)}$ 保持不变, 因为在 H_0 下, 没有 PU 信号存在。而平均检测概率 $\bar{P}_d^{(i)}$ 可以表示为^[5]

$$\bar{P}_d^{(i)} = \int_{\rho} Q_m(\sqrt{2\rho_i}, \sqrt{\lambda}) f(\rho) d\rho = e^{-\frac{\lambda}{2}} \sum_{k=0}^{m-2} \frac{1}{k!} \left(\frac{\lambda}{2}\right)^k \left(\frac{\rho}{\rho + \bar{\rho}}\right)^{m-1} \times \left(e^{-\frac{\lambda}{2(1+\bar{\rho})}} - e^{-\frac{\lambda}{2}} \sum_{k=0}^{m-2} \frac{1}{k!} \left(\frac{\lambda \bar{\rho}}{2(1+\bar{\rho})}\right)^k \right) \quad (3)$$

其中, $f(\rho)$ 是信噪比在瑞利信道下的概率密度函数, 它服从均值为 $\bar{\rho}$ 的指数分布。
假设 N 个认知用户通过独立同分布 (i.i.d.) 的衰落信道并具有相同的平均信噪比 $\bar{\rho}$ 。因此, $P_m^{(i)} = P_m$ 和 $P_f^{(i)} = P_f$, ($i = 1, 2, \dots, N$)。那么, 当使用“或”逻辑合并时, 协作频谱感知的检测性能可以表示为^[6]

$$Q_f = 1 - (1 - P_f)^N \quad Q_d = 1 - (1 - P_d)^N \quad \text{或} \quad Q_m = 1 - Q_d \quad (4)$$

在以上的分析中, 假设报告信道是无误传输的。然而, 实际中报告信道往往受到衰落的影响, 这将恶化频谱感知的性能。令 $P_{e,i}$ 表示由第 i 个报告信道所引起的误码率。假设所有的报告信道是 i.i.d., 即 $P_{e,i} = P_e$ 。

对于给定瞬时信噪比 γ , BPSK 信号在 AWGN 信道上的误码率可以表示为:

$$P_e = Q(\sqrt{2\gamma}) \quad (5)$$

因此, 瑞利信道中的平均误码率 \bar{P}_e 是将 AWGN 信道中的误码率按衰落分布进行积分

$$\bar{P}_e = \int_{\gamma} Q(\sqrt{2\gamma}) f(\gamma) d\gamma = \frac{1}{2} \left[1 - \sqrt{\bar{\gamma} / (1 + \bar{\gamma})} \right] \quad (6)$$

其中, $f(\gamma)$ 是瑞利信道下信噪比的概率密度函数, 它服从均值为 $\bar{\gamma}$ 的指数分布。

只有在两种情况下，经过报告信道译码后的本地检测结果为 H_0 ：a). 认知用户发送 H_0 ，同时通过报告信道传输后的译码信号仍为 H_0 ；b). 认知用户发送 H_1 然而译码为 H_0 由于报告信道引进的误码。因此，当考虑由报告信道的衰落特性产生的误码时，协作频谱感知的检测概率 Q_d 和虚警概率 Q_f 可以写成

$$Q_f = 1 - \left[(1 - P_f)(1 - \bar{P}_e) + P_f \bar{P}_e \right]^N \quad \text{和} \quad Q_d = 1 - \left[(1 - \bar{P}_d)(1 - \bar{P}_e) + \bar{P}_d \bar{P}_e \right]^N \quad (7)$$

在以下部分中，我们主要是通过降低 \bar{P}_e 来提高协作频谱感知的检测性能。

2 基于信道可靠性算法的性能分析

不同 SU 的检测结果有着不同的可靠性由于它的本地检测结果通过不同的信道，因此在 AP 处是有必要考虑来自各 SU 本地检测结果的可靠性。然而，在 AP 处收到的来自于具有较好报告信道 SU 的信号具有更好的可靠性。因此我们提出一种基于报告信道可靠性的协作频谱感知算法。

该算法的关键所在就是选择合适的 SUs 来参与协作感知。令在某个区域中所有可能参与协作频谱感知 SUs 的集合称之为总集合，而协作集合 $R(s)$ 是总集合的子集，它由在某个特定时间所有参与协作的认知用户构成。认知用户是否发送自己的本地二进制判决结果到 AP 参与数据的融合取决于 SU 到 AP 链路的瞬时信道增益。也就是说，只有当第 i 个认知用户到 AP 之间的链路瞬时信道增益 γ_i 大于信噪比门限值 γ_T 时，则第 i 个认知用户 s_i 满足 $s_i \in R(s)$ 。否则，该认知用户不发送它的本地二进制判决结果到 AP 参与数据的融合。我们假设至少有一个认知用户满足 $\gamma_i \geq \gamma_T$ 。

假设所有的报告信道是独立慢变化的平坦瑞利信道。同时也假定认知用户已知报告信道边信息 (channel side information, CSI)。为了测得该 CSI，AP 可以利用导频测量信道，再将测量结果反馈给 SU。

根据以上的分析可知，瑞利信道下，BPSK 在提出的基于信道可靠性的协作频谱感知算法中的平均误码率可以表示为：

$$\bar{P}_{e,RC} = \int_{\gamma_T}^{\infty} Q(\sqrt{2\gamma}) f(\gamma) d\gamma \quad (8)$$

对于给定信噪比门限值 γ_T ，第 i 个用户 s_i 成为协作集合 $R(s)$ 元素的概率为

$$Pr[s_i \in R(s)] = \int_{\gamma_T}^{\infty} \frac{1}{\gamma} \exp\left(-\frac{\gamma}{\gamma}\right) d\gamma = \exp\left(-\frac{\gamma_T}{\gamma}\right) \quad (9)$$

对于有 k 个协作用户的协作频谱感知的检测概率 $P_{d,RC}$ 和虚警概率 $P_{f,RC}$ 分别可以表示为

$$P_{d,RC} = Pr[E > \lambda | H_1, R(s)] = 1 - \left[(1 - \bar{P}_d)(1 - \bar{P}_{e,RC}) + \bar{P}_d \bar{P}_{e,RC} \right]^k \quad (10)$$

$$P_{f,RC} = Pr[E > \lambda | H_0, R(s)] = 1 - \left[(1 - P_f)(1 - \bar{P}_{e,RC}) + P_f \bar{P}_{e,RC} \right]^k \quad (11)$$

由二项概率公式，我们可以得到

$$Pr[R(s) = k] = \binom{N}{k} \left(e^{-\frac{\gamma_T}{\gamma}} \right)^k \left(1 - e^{-\frac{\gamma_T}{\gamma}} \right)^{N-k}, \quad k = 1, 2, \dots, N \quad (12)$$

其中, $|R(s)|$ 是集合 $R(s)$ 元素的个数和 $Pr[\cdot]$ 表示概率。

由全概率公式可知, 基于信道可靠性的协作频谱感知算法的平均检测概率 $\bar{Q}_{d,RC}$, 漏检概率 $Q_{m,RC}$ 和虚警概率 $Q_{f,RC}$ 分别可以表示为

$$\begin{aligned}\bar{Q}_{d,RC} &= \sum_{R(s)} Pr[E > \lambda | H_1, R(s)] Pr[R(s)] = \sum_{k=1}^N \left\{ 1 - \left[(1 - \bar{P}_d)(1 - \bar{P}_{e,RC}) + \bar{P}_d \bar{P}_{e,RC} \right]^k \right\} Pr[R(s)] \\ &= 1 - \beta - \sum_{k=1}^N \left[(1 - \bar{P}_d)(1 - \bar{P}_{e,RC}) + \bar{P}_d \bar{P}_{e,RC} \right]^k Pr[R(s)]\end{aligned}\quad (13)$$

$$Q_{m,RC} = \sum_{k=1}^N \left[(1 - \bar{P}_d)(1 - \bar{P}_{e,RC}) + \bar{P}_d \bar{P}_{e,RC} \right]^k Pr[R(s)] = 1 - \bar{Q}_{d,RC} - \beta \quad (14)$$

其中 $\beta = \left(1 - e^{-\gamma_T/\bar{\gamma}}\right)^N$ 。

$$Q_{f,RC} = \sum_{R(s)} Pr[E > \lambda | H_0, R(s)] Pr[R(s)] = \sum_{k=1}^N \left\{ 1 - \left[(1 - P_f)(1 - \bar{P}_{e,RC}) + P_f \bar{P}_{e,RC} \right]^k \right\} Pr[R(s)] \quad (15)$$

把方程(2), (3)和(8)代入方程(13), (14)和(15), 得到基于信道可靠性的协作频谱感知算法的检测性能。

3 数值仿真

为了分析提出的协作频谱感知方法的检测性能, 仿真中选取时间带宽积 m 为 5, 感知信道 (即 SU 到 PU 之间的链路) 的平均信噪比 $\bar{\rho}$ 为 10 dB, 报告信道的平均信噪比 $\bar{\gamma}$ 为 8 dB。

图 2 为协作频谱感知分别通过两种不同的报告信道下的虚警概率与漏检概率关系曲线图。依图可知, 对于固定的 N 值, 相比于理想的报告信道, 由于瑞利衰落型报告信道引进的误码, 协作频谱感知的检测性能明显的降低。此外, 无论选取哪种类型的报告信道, 检测性能都会随着 N 的增加而得到改善。此外, 从图中也可以发现, 虚警概率是有界限的。这是因为, 从方程 (7) 中, 虚警概率的下限值为 \bar{Q}_f

$$Q_f \geq \bar{Q}_f \square \lim_{P_f \rightarrow 0} Q_f = 1 - (1 - \bar{P}_e)^N \approx N \bar{P}_e \quad (16)$$

当 Q_f 减小到 \bar{Q}_f 时, 漏检概率 Q_m 会陡然增加到 1, 等价于检测概率 Q_d 很快减小到 0。这意味着认知用户不能检测到主用户的存在当 Q_f 小于 \bar{Q}_f 时。而且, \bar{Q}_f 随着 N 的增加而变大, 这正好与方程 (16) 相一致。

图 3 为信噪比门限值 γ_T 与误码率 P_e 的关系曲线图。由图中曲线变化可知, BPSK 的平均误码率 P_e 随着信噪比门限值 γ_T 的增加而减小。当 γ_T 固定时, P_e 的值随着报告信道平均信噪比 $\bar{\gamma}$ 的增加而变小。由于 γ_T 的值应当是小于 $\bar{\gamma}$, 因此, 我们可以令 $\gamma_T=5$ dB 在以下的仿真中。

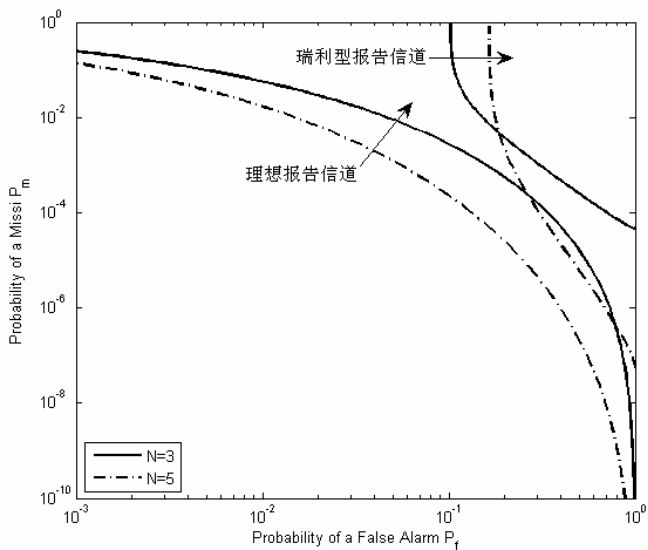


图2 传统协作频谱感知的虚警概率与漏检概率的关系曲线图
分别通过理想报告信道和瑞利衰落型报告信道

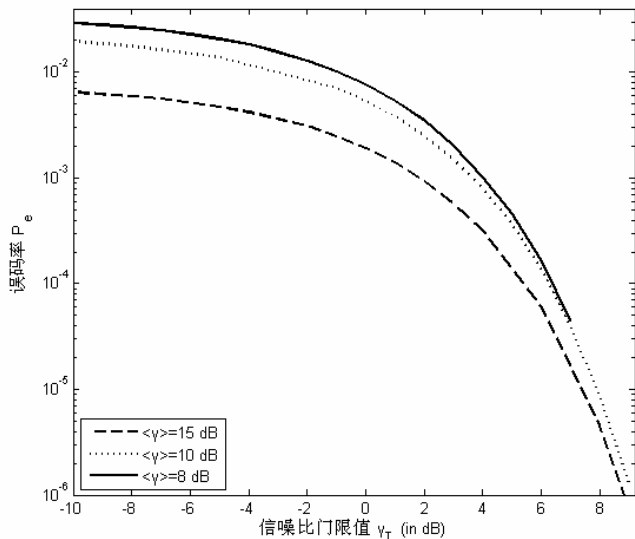


图3 BPSK 的平均误码率和信噪比门限值的关系, 报告信道的
平均信噪比分别为 8 dB, 10 dB 和 15 dB

图4表明提出的基于信道可靠性的协作频谱感知的检测性能会随着 N 的增加而变好, 这意味着 N 会产生空间分集增益。当有3个认知用户协作时, 对于给定1%的虚警概率, 对应的检测概率 P_d 大约为85%, 漏检概率 P_m 大约为13%。之所以 P_d 与 P_m 的和小于1是由于在方程(14)中 β 的存在。此外, 从图4(b)也可以发现, 虚警概率是有界限的, 这是因为, 从方程(15)中可得, $\bar{Q}_{f,RC}$ 为,

$$Q_{f,RC} \geq \bar{Q}_{f,RC} \square \lim_{P_f \rightarrow 0} Q_{f,RC} = \sum_{k=1}^N \left\{ 1 - \left[(1 - P_{e,RC}) \right]^k \right\} Pr[R(s)] \approx \sum_{k=1}^N k P_{e,RC} Pr[R(s)] = N P_{e,RC} \exp\left(-\frac{\gamma_{Th}}{\gamma}\right) \quad (17)$$

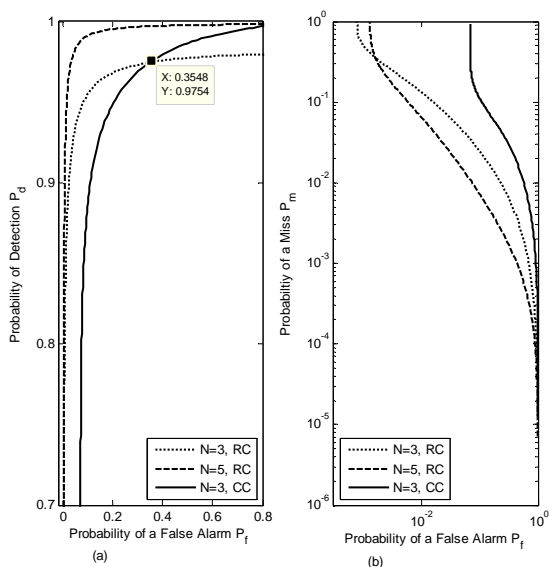


图4 基于信道可靠性的协作频谱感知算法(RC)的检测性能, $N=3$ 和 5

而且,从图中也可以发现 $\bar{Q}_{f,RC}$ 会随着 N 的增加而变大,这正好与方程(17)相一致。当 $Q_{f,RC}$ 减小到 $\bar{Q}_{f,RC}$ 时,漏检概率 $Q_{m,RC}$ 会陡然增加到 $1-\beta$,等价于检测概率 $Q_{d,RC}$ 很快的降到0。这意味着当 $Q_{f,RC}$ 小于 $\bar{Q}_{f,RC}$ 时,认知用户不能检测到主用户的存在。为了比较,我们在图4中也给出了传统协作频谱感知(CC)在 $N=3$ 时的检测性能。依图可知,当虚警概率的值较大时(约 $> 35\%$),相比于提出的基于信道可靠性的协作频谱感知算法,传统的协作方法对主用户的检测性能要更好,但如此高的虚警概率在实际中是不实用的。此外,传统协作方法中的虚警概率的下限值 \bar{Q}_f 明显的要大于提出的自适应协作频谱感知算法的 $\bar{Q}_{f,RC}$,因此,我们提出的基于信道可靠性的协作频谱感知算法更具有实用性。

4 结语

针对在协作频谱感知中由于衰落性报告信道引进的误码概率,提出了一种基于报告信道可靠性的协作频谱感知算法。通过分析和仿真的结果可以看到,相比于传统的协作频谱感知,提出的这种协作算法可以提高对主用户的检测性能,同时也更具有实用性。

参考文献

- [1] Mitola J, Maguire G Q. Cognitive radio: making software radios more personal [J]. *IEEE Pers. Commun.*, Aug. 1999, 6(4): 13-18.
- [2] Akyildiz I F, Lee W Y, and Vuran M C, et.al. NeXt generation/dynamic spectrum access/cognitive radio wireless networks: A survey [J]. *ELSEVIER Computer Networks*, 2006, 50 (13): 2127-2159.
- [3] Unnikrishnan J and Veeravalli V V. Cooperative sensing for primary detection in cognitive radio [J]. *IEEE Journal of Selected Topics in Signal Processing*, Feb. 2008, 2(1): 18-25.

- [4] Urkowitz H. Energy detection of unknown deterministic signals [J]. *Proc. of the IEEE*, April, 1967, 55(4): 523-531.
- [5] Digham F F, Alouini M and Simon M K. On the energy detection of unknown signals over fading channels [J]. *IEEE Trans. on Commun.*, Jan. 2007, 55(1): 21-24.
- [6] Liang Ying-Chang, Zeng Yonghong and Peh Edward C Y et. al. Sensing-throughput tradeoff for cognitive radio networks [J]. *IEEE Trans. on Wireless Commun.*, April 2008, 7(4): 1326-1337.

作者简介:

岳文静 (1982-), 女, 山西朔州人。上海交通大学电子工程系信息与通信工程专业博士研究生, 目前研究方向为动态频谱接入。电话: 025-83530501. E-mail: yuewenjing@sjtu.edu.cn

郑宝玉 (1945-), 男, 福建闽侯人。南京邮电大学信号处理与传输研究院教授, 博士生导师, 上海交通大学兼职教授、博士生导师。主要研究方向为现代通信中的智能信号与信息处理、认知无线电等。E-mail: zby@njupt.edu.cn

认知无线网络中基于博弈论的功率控制算法

王正强 蒋铃鸽 何 晨

(上海交通大学, 电子工程系, 上海 200240)

摘 要: 采用博弈论方法对受干扰温度约束的认知无线网络进行建模分析, 证明了该博弈问题纳什均衡解的存在性和惟一性, 并提出一个分布式的功率控制算法。该算法只需要各次用户知道局部信道状态信息, 从而比集中式的功率控制算法具有更小的信令开销。仿真结果表明该算法的具有很快的收敛速度, 同时保证了测量点处的干扰温度要求, 从而实现了对于主用户网络的严格保护。

关键词: 认知无线网络; 功率控制; 博弈论; 干扰温度

A Game Theoretic Approach to Power Control for Cognitive Network

Wang Zheng qiang Jiang Ling ge He Chen

(Department of Electronic Engineering, Shanghai Jiao Tong University, Shanghai 200240, China)

Abstract: A game theoretic solution for power control in cognitive radio network under the interference temperature constraints was proposed. The existence and uniqueness nash equilibrium of power control game is proved and a distributed algorithm is proposed. The algorithm has the advantage of a low overhead for information exchange than centralized power control algorithm since the cognitive users only need local information. Simulation results show that the convergence of the proposed algorithm is fast and interference temperature is not exceed, so the strict protection of the primary user's network is realized by the algorithm.

Keywords: cognitive network; power control; game theory; interference temperature

1 引言

认知无线网络是一个能够感知周围频谱环境的智能无线通信系统。它能够根据环境的变化自适应地调整发射功率、载波频率、编码策略等参数以提高自身网络性能和避免和降低对主用户网络的干扰^[1]。在多用户共存的认知网络中, 有效的功率控制不但可以降低认知用户间的干扰, 同时可以降低对主用户的干扰。最近, 美国联邦通信委员会 FCC(Federal Communications Commission)提出基于干扰温度的分配提案, 即当次用户网络对主用户网络的干扰小于一定门限情况下, 次用户网络可以与主用户网络共存使用^[2], 该提案增加了认知网络的接入机会, 有利于高效利用频谱。

博弈论是数学的一个重要分支。近年来,它被广泛应用到了经济学,计算机科学,通信学等众多学科领域。目前基于博弈论的功率分配和控制方面已经有了一些研究成果^{[3][4]}。但在这些文献中并没有考虑到干扰温度的影响,并且各个用户的策略调整是相互独立,没有共有约束的影响。文献[5]考虑干扰温度的影响,采用集中式优化方法,但该算法需要知道整个认知网络全部信道状态信息,同时采用复杂度较高的几何规划算法,这对于认知网络而言具有较大通信开销和计算复杂度。本文首先建立了认知无线网络的系统模型,通过应用博弈理论证明了在干扰温度约束条件下该博弈模型问题均衡解的存在性和惟一性,提出了分布式的功率控制算法。通过仿真,该分布式算法具有很快的收敛速度,且满足干扰温度测量点处的干扰温度要求。

本文的结构如下:第2节,介绍认知无线网络的系统模型;第3节,介绍带有耦合的功率控制博弈问题;在第4节,分析了该博弈问题均衡解的存在性和惟一性;第5节,给出分布式算法;第6节,进行算法仿真;第7节,结论。

2 认知无线网络系统模型

考虑如图1所示:主次用户网络共存的情形, N 个认知用户接入一个基站。同时在主用户网络周围部署 M 个测量点,用来测量次用户网络对于主用户网络的干扰。第 i 个次用户在基站的 SINR 可以表示为:

$$\gamma_i = \frac{h_i p_i}{\sum_{j=1, j \neq i}^N h_j p_j + Q + \sigma^2} \tag{1}$$

其中 p_i 是第 i 个认知用户的发射功率, h_i 是它到基站的链路增益, Q 为主用户对次用户网络的干扰, σ^2 为背景噪声。为了保证主用户网络的正常通信,避免次用户网络对于主用网络造成过大的干扰,各测量节点将监测认知网络对于主用户网络的干扰,同时反馈给认知网络,使得认知网络能够及时调整自身发射参数,来避免对于主用户网络产生较大干扰。

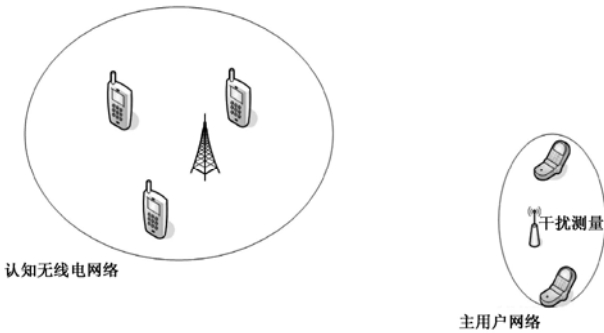


图 1

设第 m 处测量点的接收到来自次用户网络的总功率限制(干扰温度)为 $B_m(m=1,2,\cdots,M)$,即有如下 M 个不等式约束:

$$\sum_{i=1}^N p_i h_i^m \leq B_m \quad (m=1,2,\cdots,M) \tag{2}$$

其中 h_i^m 为第 i 个认知用户到第 m 处测量点的链路增益, B_m 为第 m 处测量点的干扰温度。

3 带耦合的功率控制博弈问题

对于一个博弈问题 G , 设 N 为参与博弈的参与者, Ω_i 为参与人可能采取的策略集合, U_i 为参与人在某种策略组合所获得的效用, 那么该博弈问题可以表示成:
 $G = \{N, \{\Omega_i\}_{i \in N}, \{U_i\}\}$.

定义 1: 当博弈参与者各自选择自己策略 $\bar{s}_i \in \Omega_i$, 后形成一种的状态 $\bar{s} = (\bar{s}_1, \bar{s}_2, \dots, \bar{s}_n)$, 在该状态下没有参与者愿意单方面改变自己的策略, 即: 对于每个参与人 i 和任意策略 $s_i \in \Omega_i$, 有:

$$U_i(\bar{s}_1, \bar{s}_2, \dots, \bar{s}_n) \geq U_i(\bar{s}_1, \dots, \bar{s}_{i-1}, s_i, \bar{s}_{i+1}, \dots, \bar{s}_n),$$

那么 $\bar{s} = (\bar{s}_1, \bar{s}_2, \dots, \bar{s}_n)$ 叫该博弈问题的 Nash 均衡解。

定义 2: 对于博弈问题 $G = \{N, \{\Omega_i\}_{i \in N}, \{U_i\}\}$, 设 y_{-i} 是除去用户 i 外其他用户的策略集合。若对任意参与者 i , 存在函数 P , 使得

$$U_i(x, y_{-i}) - U_i(z, y_{-i}) = P(x, y_{-i}) - P(z, y_{-i})$$

成立, 那么该博弈称为精确潜博弈, P 作为博弈问题 G 的精确潜函数^[6]。

带干扰温度约束的认知无线网络功率控制问题可以建模成如下带耦合可行策略的博弈问题:

参与者: $i \in \{1, 2, \dots, N\}$, N 个认知用户;

策略集合: 设第 i 个认知用户的功率调整范围为: $S_i = [0, p_i^{\max}]$; $S = S_1 \times S_2 \times \dots \times S_N$; N 个认知用户的可行策略集合为:

$$\Omega = S \cap \{(p_1, \dots, p_n) \mid \sum_{i=1}^N p_i h_i^m \leq B_m, m=1, \dots, M\}$$

记 $p_{-i} = (p_1, \dots, p_{i-1}, p_{i+1}, \dots, p_n)$ 为

除去第 i 个认知用户外其余认知用户的功率。那么第 i 个次用户的策略空间可以表示为:

$$\Omega_i(p_{-i}) = \{p_i \in S_i \mid \sum_{i=1}^N p_i h_i^m \leq B_m, m=1, \dots, M\},$$

从 $\Omega_i(p_{-i})$ 的定义知所有认知用户的策略空间具有耦合性。

效用函数 U_i : 表征了认知用户 i 对某种策略的偏好。我们采用如下效用函数:

$$U_i(p_i, p_{-i}) = \log(1 + \gamma) - c_i p_i^2$$

效用函数由两部分组成, 第一部分表征了认知用户的收益: 最大化自己的速率; 第二部分表征了认知用户的代价, 由于在分布式网络中, 若只考虑用户的自私性, 在多用户竞争情形下, 将降低系统性能, 其中 c_i 为代价因子。各个次用户的考虑目标为:

$$\max_{p_i \in \Omega_i(p_{-i})} u_i(p_i, p_{-i})$$

4 均衡解的存在性和惟一性

命题: 对于带耦合的功率控制博弈:

$G = \{N, \{\Omega_i(p_{-i})\}_{i \in N}, \{U_i\}\}$ 是一个精确潜博弈，它的一个精确潜函数为：

$$P(p_1, p_2, \dots, p_n) = \log\left(\sum_{i=1}^N h_i p_i + Q + \sigma^2\right) - \sum_{i=1}^N c_i p_i^2$$

证明：对于任意 i , $p_i, \bar{p}_i \in \Omega_i$,

$$\begin{aligned} & U_i(p_i, p_{-i}) - U_i(\bar{p}_i, p_{-i}) \\ &= \log\left(1 + \frac{h_i p_i}{\sum_{j \neq i}^N h_j p_j + Q + \sigma^2}\right) - c_i p_i^2 - \log\left(1 + \frac{h_i \bar{p}_i}{\sum_{j \neq i}^N h_j p_j + Q + \sigma^2}\right) + c_i \bar{p}_i^2 \\ &= \log\left(\frac{1 + \frac{h_i p_i}{\sum_{j \neq i}^N h_j p_j + Q + \sigma^2}}{1 + \frac{h_i \bar{p}_i}{\sum_{j \neq i}^N h_j p_j + Q + \sigma^2}}\right) - c_i p_i^2 + c_i \bar{p}_i^2 \\ &= \log\left(\frac{\sum_{j \neq i}^N h_j p_j + Q + \sigma^2 + h_i p_i}{\sum_{j \neq i}^N h_j p_j + Q + \sigma^2 + h_i \bar{p}_i}\right) - c_i p_i^2 + c_i \bar{p}_i^2 \\ &= \log\left(\sum_{i=1}^N h_i p_i + Q + \sigma^2\right) - \log\left(\sum_{j \neq i}^N h_j p_j + h_i \bar{p}_i + Q + \sigma^2\right) - c_i p_i^2 \\ &\quad - \sum_{j \neq i}^N c_j p_j^2 + c_i \bar{p}_i^2 + \sum_{j \neq i}^N c_j p_j^2 \\ &= P(p_1, \dots, p_i, \dots, p_n) - P(p_1, \dots, \bar{p}_i, \dots, p_n) \end{aligned}$$

从而由定义 2 知命题成立。

定理： 当对于 $\forall i \in \{1, 2, \dots, N\}$, c_i 满足：

$$c_i > \frac{h_i \left(\sum_{j \neq i}^N h_j - h_i \right)}{2 \left(\sum_{i=1}^N h_i p_i + Q + \sigma^2 \right)^2} \quad (3)$$

那么 $G = \{N, \{\Omega_i(p_{-i})\}_{i \in N}, \{U_i\}\}$ 存在惟一的 Nash 均衡。

证明：由文献[7]知只需验证两个条件：(1) Ω 是凸集, (2) 潜函数 P 是严格凹函数。

(1) 由于 Ω 是多个凸集交，故 Ω 是凸集。

$$(2) \quad \frac{\partial P}{\partial p_i} = \frac{h_i}{\sum_{i=1}^N h_i p_i + Q + \sigma^2} - 2c_i p_i$$

$$\frac{\partial^2 P}{\partial p_i^2} = -\frac{h_i^2}{\left(\sum_{i=1}^N h_i p_i + Q + \sigma^2\right)^2} - 2c_i \quad (4)$$

$$\frac{\partial^2 P}{\partial p_i \partial p_j} = -\frac{h_i h_j}{\left(\sum_{i=1}^N h_i p_i + Q + \sigma^2\right)^2}, \quad i \neq j, \quad (5)$$

由(4), (5)知, 当对任意 $\forall i \in \{1, 2, \dots, N\}$, c_i 满足 (3) 时, P 的 Hessian 矩阵是严格对角占优的, 从而 P 是严格凹函数, 从而定理得证。

5 分布式算法

将次用户基站测量的总功率记作 I :

$$I = \sum_{i=1}^N (h_i p_i + Q + \sigma^2);$$

第 i 个测量点测量的干扰记作 I_i :

$$I_i = \sum_{i=1}^N p_i h_i^m, \quad i \in \{1, 2, \dots, M\};$$

次用户采用文献[8]的最佳反应决策和文献[9]Gauss-Seidel 迭代算法依次更新各自功率, 直到各用户的状态达到稳定, 具体算法流程如下:

(1) $t=0$ 时刻各次用户初始化各自的功率:

$$p_i^0, i \in \{1, 2, \dots, N\};$$

(2) 基站和各测量点分别反馈测量的功率

$$I(t) = \sum_{i=1}^N (h_i p_i^t + Q) \text{ 和 } I_i(t) = \sum_{i=1}^N p_i^t h_i^m, \quad i \in \{1, 2, \dots, M\};$$

(3) 分布式功率调整:

$$\text{记: } \Pi_i^t = \left\{ p_i^t \mid 0 \leq p_i^{t+1} \leq \frac{B_i - I_i(t) + h_i p_i^t}{h_i} \right\} \quad \Omega_i^t(p_{-i}^t) = S_i \cap \Pi_i^t;$$

次用户按照如下方式调整自己功率: $p_i^{t+1} = \arg \max_{p_i \in \Omega_i^t(p_{-i}^t)} U(p_i, p_{-i}^t);$

重复步骤(2),(3)直到所有用户的效用函数收敛。

第(3)步的功率调整中各个次用户只需要知道自己到基站和测量点的信道增益, 不需要知道其他次用户的信道增益, 从而比集中式功率控制减少了信令开销。

6 算法仿真

考虑次用户网络包含 6 个接入用户的情形, 它们随机分布在基站周围 300m 的区域内且到基站的距离 $d = [100, 120, 140, 160, 180, 200]m$; 各个次用户功率调整范围均为 $[0, 1]W$; 两个测量点随机分布在距离次用户基站 500m 处, 干扰温度为 $10^{-10}W$; 背景噪声

$\sigma^2 = 5 \times 10^{-15} \text{ W}$ ；两个活动的主用户在测量点处以固定功率 1W 发射；干扰温度各个用户的道路径增益均为 $h = 0.097d^4$ ；价格因子均为 $c=100$ ；初始化时刻，各用户随机选择一个初始发射功率。图 2 显示了不同次用户的功率随着迭代次数的变化情况，可见各个用户经过 5 次功率调整就收敛了且满足干扰温度限制。图 3 给出了当用户 1 代价因子为 200，其余用户仍为 100 时；各用户的功率随迭代次数变化情况。可见代价因子增加将降低用户自己发射功率，同时使得均衡时其余用户的功率增加，代价因子的确定可以由用户针对不同信道状态预先设置。

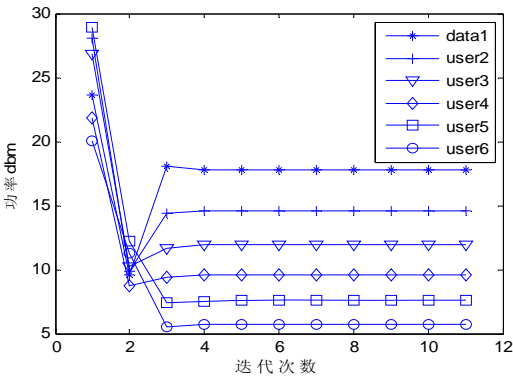


图 2

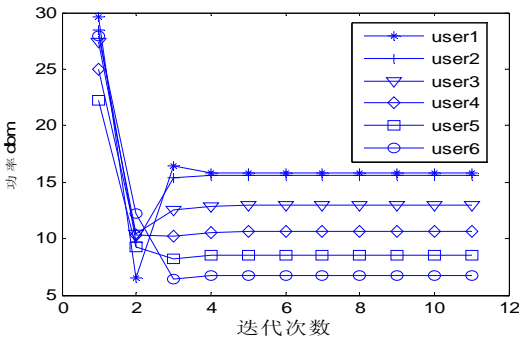


图 3

7 结论

根据主用户网络测量点处的干扰温度要求，本文采用博弈论的方法研究了次用户网络间的功率控制算法，给出了分布式算法，仿真表明该算法具有很快的收敛速度，并满足了各测量点处的干扰温度约束要求。

参 考 文 献

[1] Haykin S,“Cognitiveradio:Brain-empoweredwireless communications”, IEEE Journal on Selected Areas in Communications, vol. 23, no.2, pp. 201-220, Feb. 2005

[2] Federal Comm. Commission, Spectrum Policy Task Force, Report ET Docket no. 02-135, Nov. 2002

- [3] Goodman D, Mandayam N, “Power control for wireless data”, IEEE Personal Communications, Volume 7, Issue 2, pp. 48 – 54, Apr. 2000
- [4] Sung C W, Wong W S, “A noncooperative power control game for multirate CDMA data networks”, IEEE Transactions on Wireless Communications, Volume 2, Issue 1, pp. 186 – 194, Jan. 2003
- [5] Xing Y P, Mathur C N, Haleem M A, “Dynamic Spectrum Access with QoS and Interference Temperature Constraints”
- [6] Monderer D, Shapley L, Potential Games [J]. Games and Economic Behavior (S0899-8256), 1996, 14: 124-143
- [7] Scutari G, Barbarossa S, and Palomar D, “Potential Games: A Framework for Vector Power Control Problems With Coupled Constraints”, ICASSP, 2006.
- [8] Fudenberg D, Levine D K., The Theory of Learning in Games, MIT Press
- [9] Bertsekas D P and Tsitsiklis J N, Parallel and Distributed Computation Numerical Methods, 2nd ed. Athena Scientific Press, 1989

作者简介:

王正强: 男, 1983 年生, 四川成都人, 上海交通大学通信与信息工程专业博士生, 主要研究方向是认知无线电中资源分配。Email:wangzqiang@sjtu.edu.cn

蒋铃鸽: 女, 上海交通大学教授, 博士生导师, 主要研究方向: MIMO 系统、无线传感网、高频技术、协同通信、网络编码、无线网络的资源分配、跨层技术。Email:lgjiang@sjtu.edu.cn

何 晨: 男, 上海交通大学教授, 博士生导师, 主要研究方向: MIMO 系统、无线传感网、高频技术、协同通信、认知无线电。Email:chenhe@sjtu.edu.cn

多层网状局域连通无线传感器网络定位分析

刘 枫 郭爱煌 杭海存

(同济大学 电子与信息工程学院, 上海 200092)

摘 要: 针对地下复杂环境提出了考虑信号衰减的多层网状局域连通数学模型, 建立了多层独立的环状腔体空间。对于提出的数学模型, 使用 RSSI 和 MLE 进行定位算法的实现。通过使用迭代的方式实现了层内的节点定位, 并在层与层使用连通节点传递不同层的节点信息, 完成三维空间的定位计算, 对定位算法进行仿真。仿真结果表明: 算法能够实现节点定位, 同时, 优化后的算法与传统的基于信号强度指示测距算法相比精度有所提高。

关键词: 无线通信; 多层网状局域连通定位; 极大似然估计; 强度指示; 无线传感器网络

Localization Analysis of WSN in Multilayered Meshing Locally Connected Environment

Liu Feng Guo Aihuang Hang Haicun

(School of Electronics and Information Engineering, Tongji University, Shanghai 200092, China)

Abstract: A multilayered, meshing-locally-connected model is brought up based on the complex environment underground; a multilayer-isolated annulus cavity is built up. RSSI and MLE are used to establish the algorithm of localization. Intra-layered localization is accomplished by iterative calculating; inter-layered localization is accomplished by using connecting nodes to transfer the information in different layers. According to simulations, the results show that: localization is accomplished; meanwhile, the precision of localization is better than traditional RSSI-based algorithm. This algorithm has a widely meaning of application in the environment of underground.

Keywords: Wireless; Multilayered Meshing Locally Connected Localization; MLE; RSSI; WSN

1 引言

无线传感器网络 (Wireless Sensor Network, WSN) 的节点定位是一项复杂的技术, 由于受多径、非视距传播的影响, 使各种信号特征测量值总是存在一定偏差, 从而使各种适用于特征测量值误差服从零均值高斯分布的定位算法, 如最小二乘算法的性能显著下降, 使对目标节点的位置估计精度出现较大偏差。而且由于传感器节点资源受限的特点, 与传统的蜂窝网络和 WLAN 定位系统相比, 解决其目标节点定位问题有更大的挑战性^[1,2]。目前的算法大都在能耗、成本和精度上作了折中考虑。由于各种应用差别很大, 没有普遍适合于各种应用的定位算法, 因此要针对不同的应用, 通过综合考虑节点的规模、成本及系统对定位精度的要求, 来选择最适合的定位算法^[3]。

随着城市轨道交通的发展以及矿产资源的开发需求，地下空间的结构日趋复杂，形成多层网状局域连通，使得其环境条件与安全性的检测难度增大。在多层网状局部连通的环境中，由于节点间经常会被障碍物隔断，无需测距技术的定位算法性能明显变差^[1]。而基于测距的 WSN 定位算法相对精度较高，同时，接收信号强度指示法（Received Signal Strength Indicator, RSSI）作为基于测距的定位算法，具有通信开销小，硬件要求低的特点，在定位计算阶段，极大似然估计法（Maximum Likelihood Estimation, MLE）作为三边测量法的一种扩展，充分考虑了无线传感器网络的硬件和能耗限制而导致的误差问题，十分适合 WSN 的定位计算^[4,5,6]。基于以上出发点，本文重点分析基于 RSSI 和 MLE 的 WSN 定位算法。

2 定位算法构建分析

传感器网络的节点定位过程通常分为两个阶段。第一个阶段是测距阶段，未知节点首先测量到邻居节点的距离，然后进一步计算到临近信标节点的距离。第二个阶段是定位阶段，未知节点在计算出到达三个或三个以上信标节点的距离后，利用三边测量法、极大似然估计法或 Min-max 定位法计算未知节点的坐标。

2.1 测距阶段

基于 RSSI 的测距方法是通过测量接收机获得的信号强度，采用适当的电波传播模型对目标节点的位置进行判断的一种定位方法。该方法的关键是通过测量信号的衰减程度来估计目标节点和多个定位节点之间的距离以及目标节点之间的距离，然后再通过这些距离估计目标节点的坐标。

路径损耗 $PL(r)$ 为

$$PL(r) = 10\log(P_t) - 10\log(P_r) = PL(r_0) + 10k \log\left(\frac{r}{r_0}\right) \tag{1}$$

式(1)中 P_t 为发射功率， P_r 为接受功率， r_0 为参考距离，可以设为单位距离。由于接收信号会受到墙体和其它物体的阻碍而使得接收信号强度围绕着平均值波动，即慢衰落，为了包括这种影响，要对式(1)做出修改，增加一项随机分量 X ，它是标准差为 σ ，平均值为 0 的高斯分布随机变量^[2]。

综上所述，距离与场强之间的一般意义上的解析式为

$$r = r_0 \times 10^{\left[\frac{10\log(P_t) - 10\log(P_r) - PL(r_0) - X\sigma}{10k}\right]} \tag{2}$$

2.2 定位阶段

得到目标节点和多个定位节点之间的距离以及目标节点之间的距离之后，就可以进入无线传感器网络定位算法的定位阶段。三边测量法、极大似然估计法、Min-max 定位法是目前较常用的三种定位法，本文使用的是极大似然估计法。其基本原理就是寻找一个使测距距离与估算距离之间差异最小的点，以该点为未知节点的位置。假设 0 为未知节点，坐标是 (x_0, y_0) ，而 1、2、3、4 均为定位节点，坐标分别为 (x_1, y_1) 、 (x_2, y_2) 、 (x_3, y_3) 、 (x_4, y_4) ，0 节点到四个定位节点的距离分别是 $\rho_1, \rho_2, \rho_3, \rho_4$ 。设

$$f_{i,0} = \rho^2 - ((x_i - x_0)^2 + (y_i - y_0)^2) \quad (3)$$

式中, ρ_i 表示节点 i 与节点 0 的测量距离, 而 $((x_i - x_0)^2 + (y_i - y_0)^2)$ 则表示节点 i 与节点 0 的估算距离的平方。求节点 0 坐标 (x_0, y_0) , 使得 $\sum f_{i,0}$ 之和最小。

在无线传感器网络中, 节点之间的连接很多, 大多数节点都可以直接或间接的获得三个以上参考点的不精确距离以在二维空间中确定一个点的坐标。在文献[4]中, 作者收集的数据显示了在距离测量误差较大的情况下 (超过标准 50%), 在二维空间内定位一个被 5 个或更多已知位置节点包围的节点时, 所获得的定位位置与实际的平均偏差小于射频距离的 5%。

3 模型构建与仿真分析

3.1 仿真模型构建

地下空间是特殊而具有重要研究意义的环境场合。由于地层和岩层对电磁波的损耗阻碍了其传播效果, 造成了地下空间的局域性, 即可视空间 (高频电磁波传输空间) 是局部的、环状连接成网的、多层结构的, 基于这种环境的无线传感器网络既有平面局部联通分布, 也有垂向分层分布, 形成多层网状联通的空间分布。基于上述环境特点, 抽象出如图 1 所示的空间模型。

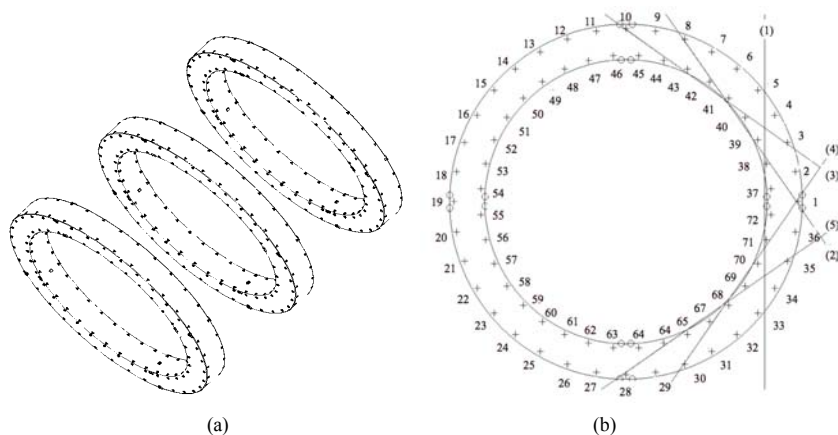


图 1 多层网状局域联通环境模型示意图

其中部署了三类传感器网络节点:

(1) 普通节点: 按照一定的规律布置在环境的墙壁、地板和天花板上。负责检测环境中的事件, 并通过多跳路由的方式将感知信息发送出去。

(2) 联通节点: 在图 1 中的三层环境中每一层的内侧均匀放置四个连通节点。其通信能力、计算能力和电池寿命都要远强于普通节点和定位节点, 其主要功能是负责各层之间的信息互通, 保证整个环境的信息一致性。

(3) 定位节点: 如图 1 所示在每个平面每隔 90 度放置四个此类节点。这些节点是通过人工部署来获得完整的坐标信息, 用来作为普通节点定位时的坐标参考依据。

3.2 特定模型下定位算法的改进与实施

本文采用迭代计算和加权平均相结合的方法，使得定位节点和未知节点的数目比例控制在一个较小的范围之内的同时，未知节点的定位可以达到比较高的精度。

在图 1(b)中，首先，三点钟方向上的四个定位节点共同的视线距离范围是以线(1)为界以右的所有的未知节点，即 1-4, 34-38, 71 和 72 号节点，这 11 个未知节点即可以通过四个定位节点在迭代计算的第一轮定位得出。得到上述 11 个节点的坐标信息之后，这 11 个在第一轮计算中作为未知节点的节点在之后的计算中就可以充当定位节点，通过这些“伪定位节点”来定位其它的未知节点。在迭代计算的第二轮，通过调用编号为 1、2、3、4 的四个在迭代计算第一轮中获得位置信息的节点的坐标，可以定位出线(2)右上的这些位置节点，包括 5-8、39-41 号节点。同样，这些未知节点在迭代计算的下一轮也变成了“伪定位节点”提供定位计算的基准。以此类推，通过这样不断的将未知节点的坐标进行定位计算而使其在迭代计算的下一轮变成定位计算的“伪定位节点”，理论上可以仅需要一组定位节点就可以将所有的未知节点坐标信息通过迭代计算得出。但由于每一轮的计算都会产生误差，而迭代计算会产生误差积累，仿真表明，其定位误差随着每一轮计算而呈近似等比增长，如图 2(a)所示，使得迭代计算在距离定位节点三轮之后的定位结果变得没有参考价值，于是综合考虑误差和定位节点占有节点的比重两方面考虑，确定放置四组定位节点在如图 1(b)所示的四个方位上，用“o”符号表示，每组节点负责除依靠本定位节点定位之外的上下各两轮迭代计算，即每组节点计算五轮迭代。图 2(b)则给出了根据这种原则进行定位计算时迭代计算轮次与误差积累的关系示意图。

经过迭代计算之后，每个点都会有两到三个不同组定位节点发起的定位计算所测得的坐标，需要按照可靠程度的不同为这些坐标设定权值，使其加权平均之后更加接近真实的节点位置。

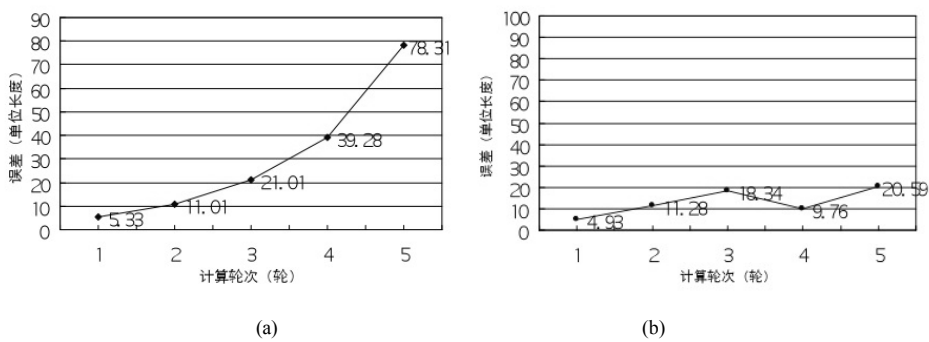


图 2 迭代轮次与误差累计关系图

由于仿真环境是多层的，单纯的二维坐标已经不能满足传感器节点定位的需求，而需要引入三维坐标。解决这一空间定位问题，可以采用 GPS 技术所使用的四点空间定位，但由于 GPS 的使用环境和精度限制，不适合在该仿真环境中应用。另外一种方法是将节点编号，并且在环境中抛撒节点时按照节点的编号，在一定编号范围内的节点放置在一定的平面上，而节点接受到不同编号范围内的节点信号时不予理会，这样就有效控制了不同 z 坐标的节点相互干扰产生的定位计算错误问题。

3.3 仿真结果分析

根据第 3.2 节的无线传感器定位改进算法，在多层网状局域连通的环境下进行仿真分析。为了提高定位精度，以下如无特殊说明，均采用 10 次定位计算之后的平均值作为定位计算的最终结果。

定位的迭代计算过程如图 3(a)所示，图中符号“o”表示“伪定位节点”，即通过定位节点在迭代计算的上一轮过程定位出的未知节点，在本轮计算中充当定位的基准。由这些“伪定位节点”定位出的节点坐标信息在图中由实心圆点表示。经过平均计算，得到最终的未知节点坐标由图中符号“+”表示，误差椭圆表示每次计算节点坐标值和最终的节点坐标值的贴近程度，并反映了误差在各个方向上的分布情况。

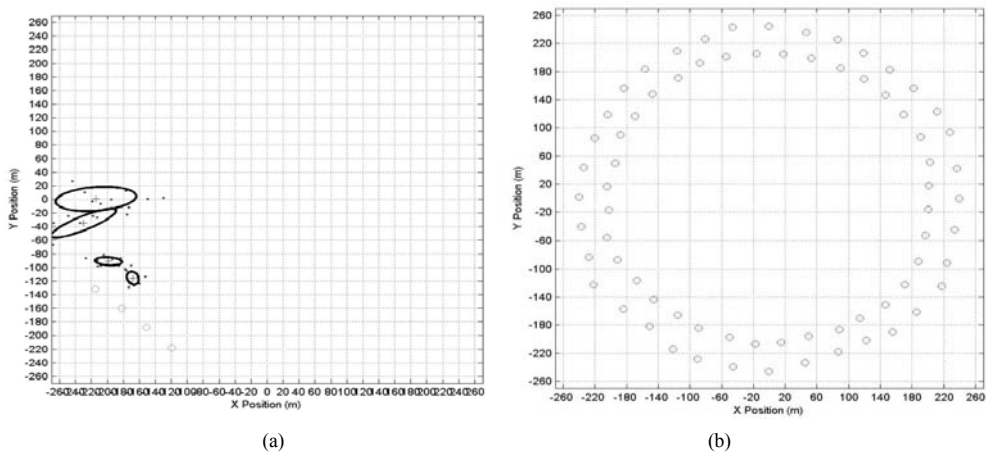


图 3 定位过程节点坐标示意图

经过完整的静态节点定位迭代计算过程之后得到每个平面上的坐标信息，此时由于每个未知节点都会有两到三个不同组定位节点发起的定位计算所测得的坐标信息，所以坐标信息的个数多于实际未知节点的个数。

由于迭代计算会造成误差累计，为了进一步减小误差，提高定位精度，将迭代计算得到的信息进行加权平均，得到图 3(b)所示的最终未知节点信息图，其中图上方的“1”表示三层的网状局域连通模型中的第一层下平面。

对于定位算法的三维扩展，根据 3.2 节的算法改进方案，将节点编号，每平面 72 个节点，并且在环境中抛撒节点时按照节点的编号，在一定编号范围内的节点放置在一定的平面上。

在定位计算的过程中必然会产生一定的误差，然而可以通过延长计算时间、增加定位次数来减小误差。误差随着定位次数的增加而减小，当定位次数大于 13 次之后，减小的趋势放缓。由于无线传感器网络是功耗敏感的，在保证定位精度的前提下，必须尽可能减小定位过程中的能量消耗，而能量的消耗和算法消耗的时间是直接相关的。定位次数的增加造成算法消耗时间的线性增长。所以权衡算法消耗时间和精度的关系，当定位次数为 13 次时，消耗的时间为 849.7ms，平均定位误差为 4.4m，精度在 10%左右，能够达到传感器网络的定位精度要求。

4 结论

通过分析地下复杂多变的环境,建立了多层网状局域连通的数学模型,根据课题的实际应用环境建立了信道传输模型并且修改了无线传输信道参数,使用接收信号强度指示法和极大似然估计法来构建无线传感器网络节点的定位算法,并根据建立的数学模型使用了迭代计算和加权平均的方法在构建的算法基础上进行优化。仿真模拟了修改后的定位算法,同时讨论了定位算法的三维扩展问题,并对定位过程中的误差与计算消耗时间的关系进行了探讨。

仿真结果表明,修改后的定位算法通过迭代计算和加权平均,有效的解决了定位节点数量要求少和定位视线距离有局限的矛盾。在对多层网状局域连通的方针环境中进行多层的空间化节点定位仿真过程中,引入编号的方法,将普通节点进行编号,这样便能够在搭建的多层网状局域连通的理想环境下实现三维空间的节点定位,同时避免了不同层节点之间的干扰。在对定位算法的次数、消耗时间和定位误差进行讨论之后,确定了定位次数为 13 次时能够在相对较少的时间内达到比较理想的定位精度,理想情况下定位精度为 10%左右。

参 考 文 献

- [1] 史龙,王福豹,段渭军等. 无线传感器网络 Range-Free 自身定位机制与算法[J]. 计算机工程与应用, 2004: 127~130, 151.
- [2] 刘剑等译. 无线通信原理与应用[M]. 北京: 清华大学出版社, 2002.
- [3] 史龙,王福豹. 无线传感器网络自身定位算法研究[M]. 西安: 西北工业大学, 2005.
- [4] Yuechen Zhang, Liang Cheng. PLACE: protocol for location and coordinate estimation: a wireless sensor network approach [J]. Computer Networks: The International Journal of Computer and Telecommunications Networking, Volume 46, Issue 5. 2004, 12.
- [5] Javid Taheri, Albert Y. Zomaya. Clustering techniques for dynamic location management in mobile computing [J]. Journal of Parallel and Distributed Computing, 2007, 4(7) :430-447 .
- [6] 姜立娣,沈建华. 基于 WSN 定位技术的研究与实现[D]. 上海: 华东师范大学, 2008

作者简介:

刘枫(1984-), 男, 辽宁人, 硕士研究生, 研究方向为无线传感技术与新一代宽带移动通信网;

郭爱煌(1964-), 男, 江西人, 博士, 教授, 研究方向无线网络与宽带通信技术;

杭海存(1985-)女, 江苏人, 硕士研究生, 研究方向为新一代宽带移动通信网与通信信号处理。

DVB-RCS中循环递归Turbo码尾比特译码技术

巩克现 刘世刚 葛临东

(中国人民解放军信息工程大学 信息工程学院, 河南 郑州 450002)

摘要: 研究了 DVB-RCS 中采用尾比特技术的循环递归 Turbo 码的译码技术, 利用此码的循环递归特性, 避免利用特征向量求解初态的译码算法, 译码起始时不用知道起始状态, 并通过仿真说明此算法比准确知道起始状态仅仅损失 0.02dB, 此算法已经在数字视频广播回传信道中采用。

关键词: DVB-RCS; Turbo 码; MAP 译码; 尾比特译码; 迭代译码

Tail-biting Decoding of Circular Recursive Systematic Turbo Code in DVB-RCS

Gong Ke xian Liu Shi gang Ge Lin dong

(Institute of Information Engineering, Information Engineering University, Zhengzhou 450002, China)

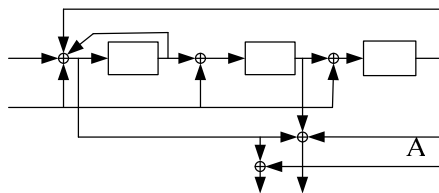
Abstract: Tail-biting decoding of circular recursive systematic Turbo code in DVB-RCS is studied and avoided. The starting state is unknown and the eigenvector is avoided when decoding. The simulation shows that there is 0.02dB loss compared with the known state. This algorithm has been used in return channel of digital video broadcast.

Keywords: DVB-RCS; Turbo Code; MAP Decoding; Tail-biting Decoding; Iterative Decoding

DVB 在卫星回传信道标准 DVB-RCS 标准^{[1][2]}中采用了多进制循环递归 Turbo 码, 此类信号已经大量出现在军事和民用当中, 标准中采用的 turbo 码的子码为卷积码, 一般地, 卷积码软输入软输出的最佳译码需要知道初始状态^{[5][6][7]}, 但标准中采用了循环递归卷积码, 初态是不固定的, 这对性能有一定的影响, 本文重点研究循环递归卷积码的译码算法与性能分析。

1 双比特循环递归系统卷积码的尾比特编码技术

k 个信息比特组成 $N = k / 2$ 个比特对进入编码器, 每个分组的比特对的个数为 $N \in \{48, 64, 212, 220, 228, 424, 432, 440, 752, 848, 856, 864\}$, 每个分组的字节数为 $N / 4$, 图 1 中 A 表示比特对中第一个比特, B 表示第二个比特, 两个校验比特用 W 和 Y 表示。



S_1

图 1 DVB-RCS 中多进制循环递归系统卷积子码编码器

定理 1: 令 $\mathbf{S}_k = [S_{k,1} S_{k,2} S_{k,3}]^T$, $S_{k,m} \in \{0,1\}$ 表示 k 时刻编码器的状态,

$$\mathbf{S}_{k+1} = \mathbf{G}\mathbf{S}_k + \mathbf{X}_k \quad (1)$$

其中

$$\mathbf{X}_k = \begin{bmatrix} A_k + B_k \\ B_k \\ B_k \end{bmatrix} \quad (2)$$

$$\mathbf{G} = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \quad (3)$$

由于码字的尾比特的特点, 每个分组必须由子编码器进行两遍编码, 第一遍编码时, 编码器初始化为全 0 状态, $\mathbf{S}_0 = [000]^T$, 这次编码的最终状态为 \mathbf{S}_N , 则以状态

$$\mathbf{S}_c = (\mathbf{I} + \mathbf{G}^N)^{-1} \mathbf{S}_N \quad (4)$$

为初态, 对同样的数据进行第二次编码, 最终状态也为 \mathbf{S}_c , 称为循环状态。

证明: 以初态 $\mathbf{S}_0 = [000]^T$ 进行编码, 则由式(1)易知, 最终状态为

$$\mathbf{S}_N = \mathbf{G}^{N-1} \mathbf{X}_0 + \mathbf{G}^{N-2} \mathbf{X}_1 + \cdots + \mathbf{X}_{N-1}$$

以 \mathbf{S}_c 为初态进行编码, 设最终状态为 \mathbf{S}'_N , 则

$$\begin{aligned} \mathbf{S}'_N &= \mathbf{G}^N \mathbf{S}'_{N-1} + \mathbf{X}_{N-1} \\ &= \mathbf{G}^N \mathbf{S}_c + \mathbf{G}^{N-1} \mathbf{X}_0 + \mathbf{G}^{N-2} \mathbf{X}_1 + \cdots + \mathbf{X}_{N-1} \\ &= \mathbf{G}^N (\mathbf{I} + \mathbf{G}^N)^{-1} \mathbf{S}_N + \mathbf{S}_N \\ &= [\mathbf{G}^N (\mathbf{I} + \mathbf{G}^N)^{-1} + \mathbf{I}] \mathbf{S}_N \\ &= [\mathbf{G}^N (\mathbf{I} + \mathbf{G}^N)^{-1} + (\mathbf{I} + \mathbf{G}^N)(\mathbf{I} + \mathbf{G}^N)^{-1}] \mathbf{S}_N \\ &= [\mathbf{G}^N + \mathbf{I} + \mathbf{G}^N] (\mathbf{I} + \mathbf{G}^N)^{-1} \mathbf{S}_N \\ &= (\mathbf{I} + \mathbf{G}^N)^{-1} \mathbf{S}_N \\ &= \mathbf{S}_c \end{aligned}$$

证毕

2 尾比特MAP译码技术

数据译码过程中 MAP 译码器最终是以概率信息或“软输出”来表示比特的可靠性，经典的 MAP 译码器是 BCJR 算法，BCJR 算法是用于网格码和 Markov 数据源的 MAP 译码器，文献[8]中给出的算法需要知道网格的起始和终止状态，或者至少有一个先验概率，通常，网格两端的某个状态的先验概率为 1，强制编码网格路径在译码终止时终止于某个状态，称为终止路径，如果不强制网格终止，译码器性能在译码器网格的结尾处会有较弱的纠错能力。

DVB-RCS 中的 Turbo 码的路径终止采用了尾比特技术，不会降低码率，也不会降低误码性能，在这种编码技术中，网格起始和终止于同一个状态，可以看成是一个循环网格，如图 2 所示，这是一个 4 状态的网格，早期全面阐释尾比特技术的是 Ma 和 Wolf^[6]。

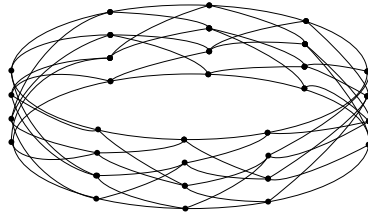


图 2 四状态循环尾比特网格

为了处理接收的网格码，接收到观测信道值后，BCJR 译码器必须要从起始和终止状态的条件分布开始，但是对于尾比特技术来说，这两个状态的分布并不知道，虽然 BCJR 译码器可以认为 M 个起始和终止状态都是均匀分布的，即 $(1/M, \dots, 1/M)$ ，但是这样做的结果同准确地知道状态相比还是要差一些，在本文中，把 BCJR 算法扩展到起始状态和终止状态未知的情况，就可以将尾比特技术应用到迭代译码的过程中[2]。

根据文献[1]，定义三个概率的集合，用矩阵和向量表示很方便，定义状态转移矩阵为 γ_k ，矩阵中的元素为

$$\gamma_k(s', s) @ P\{S_k = s, y_k | S_{k-1} = s'\} \quad k = 1, \dots, N \quad (5)$$

其中， s', s 取遍 M 个状态，将前向状态概率写成行向量 α_k 的形式，其中的元素为

$$\alpha_k(s) @ P\{S_k = s, y_1^k\}, \quad k = 1, \dots, N \quad (6)$$

类似地定义列向量 β_k ，其中

$$\beta_k(s) @ P\{y_{k+1}^N | S_k = s\}, \quad k = N-1, \dots, 1 \quad (7)$$

为了推导方便，引入联合概率：

$$\lambda_k(s) = P\{S_k = s, y_1^N\} = \alpha_k(s) \beta_k(s) \quad (8)$$

$$\sigma_k(s', s) = P\{S_{k-1} = s', S_k = s, y_1^N\} = \alpha_{k-1}(s') \gamma_k(s', s) \beta_k(s) \quad (9)$$

函数 γ_k 是算法的输入， α_k 、 β_k 是运算的过程中计算出来的。给定集合 γ_k ，基本的 BCJR 算法的步骤如下：

1) 由前向递归算法构造行向量 $\alpha_1, \dots, \alpha_L$

$$\alpha_k = \alpha_{k-1} \gamma_k, \quad k = 1, \dots, N \quad (10)$$

2) 由后向递归算法计算列向量 $\beta_{N-1}, \dots, \beta_1$

$$\beta_k = \gamma_{k+1} \beta_{k+1}, \quad k = N-1, \dots, 1 \quad (11)$$

3) 构造行向量 $\lambda_1, \dots, \lambda_N$

$$\lambda_k = \alpha_k \mathbf{g} \beta_k, \quad k = 1, \dots, N \quad (12)$$

在 BCJR 算法的标准应用中, 前向递归初始化为 $\alpha_0 = (1, 0, \dots, 0)$, 后向递归初始化为 $\beta_N = (1, 0, \dots, 0)$, 这等价于假定编码器初始并终止于 0 状态, 注意到 α_k 的各个分量之和就是观测到的 y_1^k 的概率, 如果把它归一化, 就变成了时刻 k 的状态概率, BCJR 也计算了在时刻 k 状态从 s' 转移到 s 的概率, 即

$$\begin{aligned} \sigma_k(s' \rightarrow s) @ P\{S_{k-1} = s', S_k = s, y_1^N\} \\ = \alpha_{k-1}(s') \gamma_k(s', s) \beta_k(s) \end{aligned} \quad (13)$$

BCJR 方法假定网格上的状态变化是一个 Markov 过程, 信道噪声是无记忆的。在这个假设下, 式(5)的概率 $\gamma_k(s', s)$ 变成

$$\begin{aligned} P\{S_k = s, y_k | S_{k-1} = s'\} \\ = P\{y_k, \text{状态转移 } s' \rightarrow s\} \\ = P\{y_k | s' \rightarrow s\} P\{s' \rightarrow s\} \end{aligned} \quad (14)$$

现在来研究新的尾比特 BCJR 算法, α_0 是未知的, 必须想办法求出, 求解的算法依赖于下面的定理,

定理 2: α_0 是矩阵 $\gamma_1 \mathbf{L} \gamma_N / P\{y_1^N\}$ 特征值为 1 的左特征向量, 即

$$\alpha_0 = \frac{\alpha_0 \gamma_1 \mathbf{L} \gamma_N}{P\{y_1^N\}} \quad (15)$$

证明: 网格起始分布 α_0 , 由式(10)来迭代 N 次, 得到

$$\alpha_N = \alpha_0 \gamma_1 \mathbf{L} \gamma_N \quad (16)$$

式(16)对于任何的起始分布都是成立的, 考虑一个特定的情况,

$$\alpha_0 = P\{S_0 = s | y_1^N\} = P\{S_0 = s, y_1^N\} / P\{y_1^N\}$$

这是观测到信道输出的条件下, 状态 S_0 的分布, 由于采用尾比特技术 $S_0 = S_N$, 因此对于所有的状态 s , 有

$$P\{S_0 = s, y_1^N\} = P\{S_N = s, y_1^N\} = \alpha_N(s)$$

因此 $P\{y_1^N\} \alpha_0 = \alpha_N = \alpha_0 \gamma_1 \mathbf{L} \gamma_N$

$$\text{即 } \alpha_0 = \frac{\alpha_0 \gamma_1 \mathbf{L} \gamma_N}{P\{y_1^N\}}$$

在给出算法之前, 进一步来看一下式(15), 根据式(15), α_0 是矩阵 $\gamma_1 \mathbf{L} \gamma_N / P\{y_1^N\}$ 的特征值为 1 的左特征向量, 这样的特征值一定会存在, 因为式(15)是概率的方程, 而且不存在比 1 更大的特征值, 也就是说, α_0 是 $\gamma_1 \dots \gamma_N$ 最大的特征值 $P\{y_1^N\}$ 的特征向量, 然而, 为了求 α_0 , 并不需要精确地找到 $P\{y_1^N\}$, 这是因为随着 N 的增加 $P\{y_1^N\}$ 变得非常小^[10], 比如说, 随着乘积项 $\gamma_1 \mathbf{L} \gamma_N$ 逐渐的增大, 可以用一个合适的因子来归一化来保证矩阵的元素保持在一个合

理的大小，特征向量的方向将保持不变，最后向量可以被归一化，因为仅仅需要求出最大的特征值的特征向量，并且 γ 矩阵是非常稀疏的，这就避免了求全部特征向量^[3]，降低了复杂度。

如前所述，由 λ_k 就可以求出 $P\{S_k=s'|y_1^N\}$ ，根据式(13)就可以求出 $\sigma_t(s' \rightarrow s)$ 。

在大部分BCJR应用中， $P\{y_1^N\}$ 只是用来归一化 λ ，当 $P\{y_1^N\}$ 并不需要准确地求出后，后面的方案采用连续的归一化，控制计算的精度，这个方案基于这样的事实：在式(10)或式(11)的递归过程中， α_k 、 β_k 、 γ_k 的归一化仅仅是使得式(12)中 λ_k 的归一化。但是把 λ_k 归一化为和为 1 后， $P\{S_t=i|y_1^N\}$ 仍然是正确的。

前面重归一化的方案没有计算精确的特征向量，事实上，如果前向迭代足够的次数，并将结果适当地归一化，那么最后的结果序列 $\alpha_1, \alpha_2, L, \alpha_N, \alpha_{N+1}, L, \alpha_{2N}, \alpha_{2N+1}, L$ 将收敛到 $\alpha_1^0, \cdots, \alpha_N^0$ ，这是因为一个向量反复乘上矩阵 G 后生成的向量收敛到 $\lambda_{\max}^n v$ ，其中 λ_{\max} 是 G 最大的特征值， v 是特征向量^[3]。

一个更加充分的理由如下，定义 $G = K_0 \gamma_1 L \gamma_N$ ，其中 K_0 是归一化因子，保证 G 最大的特征值为 1，从某个状态分布 α_0 开始迭代，经过 N 次迭代之后，得到

$$\alpha_N^* = K_0 \gamma_1 L \gamma_N \alpha_0 = G \alpha_0$$

其中 * 表示 α_N 被 K_0 归一化，这将产生集合 $\alpha_1^*, L, \alpha_N^*$ ，再经过 N 次迭代，产生 $\alpha_{N+1}^*, L, \alpha_{2N}^*$ ，依次类推计算到第 n 组，（假定 $\gamma_n = \gamma_m$ ，其中 $m \equiv n \bmod N$ ）序列 $\alpha_0, \alpha_N^*, \alpha_{2N}^*, L, \alpha_{nN}^*$ 一定收敛到一个极限，经过归一化后，这个极限是 α_0 ，类似地， $\alpha_{nN}^*, L, \alpha_{nN-1}^*$ 经过归一化后收敛到 $\alpha_1^0, L, \alpha_N^0$ 。

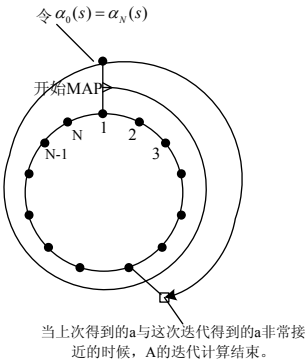


图3 尾比特 BCJR 译码器示意（所有的 α 都归一化成和为 1）

合理地选择初始化 α_0 将会导致更快的收敛，一般选择 $\alpha_0 = (1/M, L, 1/M)$ ，看上去似乎是要经过很多次 N 周期才能收敛，但事实上，前向递推很快就“忘记”了错误的起始状态估计 α_0 ，图 3 显示了前向递推如何再次环绕网格一周，在 α_{N+N} 变得接近第一圈 α_t 之前的额外的网格搜索称之为环绕深度 W ，在很多应用中，是小于 N 的，基本的迭代过程与原始的 BCJR 算法是相同的，因此算法复杂度比标准的 BCJR 算法要大 $(W + N)/N$ 。

类似地，反向递推将收敛到期望的集合 β_1^*, L, β_N^*

接收到 $y_{1,L}, y_N$ ，取 $\alpha_0 = (1/M, \dots, 1/M)$ ，新的算法如下所示：

1) 按照下面的递推式求出 α_k ，再归一化行向量 $\alpha_1^*, \dots, \alpha_L^*$ ，

$$\alpha_k = \alpha_{k-1} \gamma_k,$$

$$\alpha_k^* = \alpha_k / \sum_{s'} \alpha_k(s'), \quad k=1, L, N-1$$

继续迭代来求出 α_k^* ， $k=N+1, N+2, L$ 取 $\gamma_k = \gamma_n$ ，其中 $n \equiv k \bmod N$ ，当 $\|\alpha_k^* - \alpha_n^*\|$ 足够小时停止，此时这最后的 N 个 α 向量就得到了。

2) 用相似的方法沿着反向递推就可以得到 $\beta_1^*, \dots, \beta_N^*$ 。

3) 令 $\lambda_k = \alpha_k^* \cdot \beta_k^*$ ，

$$\lambda_k^* = \lambda_k / \sum_s \lambda_k(s), \quad k=1, L, N。$$

3 仿真结果

本部分通过仿真展示了 DVB-RCS Turbo 码的性能，仿真条件：信息对长度为 $N=212$ ，码率 $1/3$ ，几种译码算法的误比特率如图 4 所示，仿真结果表明：由于循环状态的求解不正确而导致的误码率损失大概只有 0.02dB。

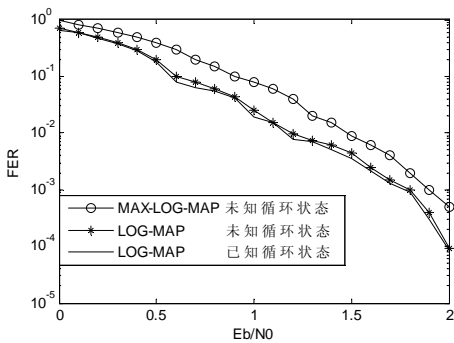


图 4 仿真结果

参 考 文 献

- [1] European Telecommunications Standards Institute. Digital broadcasting system for television, sound, and data services[M]. ETS 200 421, 1994.
- [2] European Telecommunications Standards Institute. Digital video broadcasting (DVB); interaction channel for satellite distribution systems[M]. ETSI EN 301 790V1.2.2 (2000-12), 2000.
- [3] European Telecommunications Standards Institute. Digital video broadcasting (DVB) second generation framing structure, channel coding and modulation systems for broadcasting, interactive services, news gathering and other broadband satellite applications[J]. DRAFT EN 302 307 DVBS2-74r15, 2003.
- [4] P. Robertson, P. Hoeher, and E. Villebrun. Optimal and sub-optimal maximum a posteriori algorithms suitable for turbo decoding[J]. European Trans. On Telecommun., 8(2):119-125, Mar./Apr. 1997.

- [5] C. Berrou, C. Douillard, and M. Jezequel. Multiple parallel concatenation of circular recursive convolutional (CRSC) codes[J]. Annals of Telecommunication, 54(3-4):166–172, Mar.-Apr. 1999.
- [6] H. H. Ma and J. K. Wolf. On tail biting convolutional codes[J]. IEEE Trans. Commun., 34:104–111, May 1986.
- [7] C. Berrou and M. Jezequel. Non-binary convolutional codes for turbo coding[J]. IEE Electronics Letters, 35(1): 39–40, Jan. 1999.
- [8] J. B. Anderson, New Reduced-state BCJR Algorithms[J], IEEE Trans. Commun., No.6, 2007

作者简介:

巩克现（1976-），男，山东泰安人，信息工程大学博士研究生，主要研究方向：差错控制编码技术、软件无线电、联合编码调制技术。

刘世刚（1972-），男，河南开封人，信息工程大学讲师、博士，主要研究方向：无线通信信号接收。

葛临东（1946-），男，山东济南人，教授，博士生导师，主要研究方向：软件无线电、信号分析与处理。

基于协同中继的异构无线网络切换性能分析

高智伟

(同济大学宽带无线通信与多媒体研究室, 上海 201804)

摘要: 近年来提出的协同中继技术能够有效提高网络容量, 备受科研工作者和设备厂商的大量研究和广泛关注。本文分析了未来网络的覆盖环境以及协同中继的研究状况。将协同中继模型引入至异构网络接入环境中, 研究了在此场景下, 用户切换时所能获得的容量增益和中断概率, 并推广至多用户网络中, 以验证多覆盖小区的性能。通过理论仿真分析, 证明了在异构网络环境中, 通过增加协同中继节点, 能给整个小区带来明显的性能提升。
关键字: 协同中继; 异构网络; 中断概率; 多用户分集

Cooperative Relay Based Performance Analysis of Heterogeneous Wireless Network Handover

Gao Zhi Wei

(Broadband Wireless Communication and Multimedia Laboratory, Tongji University, Shanghai 201804)

Abstract: Cooperative relay technique proposed in recent years could improve network capacity effectively, which has been widely investigated and received extensive attention by scientific researchers and equipment manufacturers. In this paper, the future network coverage environment and the research of cooperative relay are analyzed in detail. Cooperative relay model is introduced to heterogeneous wireless access environment. In this scenario, higher capacity gain and outage probability can be achieved when handover of MS takes place. And the multi-coverage cellular performance is evaluated by analyzing the multiuser diversity gain. The cellular performance improved by increasing the cooperative relay node in the heterogeneous network is proved through theoretical simulation.

Keywords: CooperativeRelay; HeterogeneousNetwork; OutageProbability; MultiuserDiversity

引言

在未来网络中, 异构性将会是一个最主要的特点, 在同一个小区将会有多种不同接入技术的无线通信网络相互交叠覆盖。对于应当采取何种方式实现异构网络的互联互通, 大体分为紧耦合与松耦合两种[1], 而对于数据网络的融合如现有的 3G 和 IEEE802.21, 业界偏向于采用松耦合的方式, 即在一个公共节点上, 提供多种网络的接入, 公共接入节点之间通过免

认证的频段或者有线的技术相连形成接入层的网状网，用以提供高速的 P2P 业务。这个公共节点可以看成是一个增强性的协同中继接入节点，在最近的 LTE - Advanced 提案中，也明确定义了 Type2 中继[2]，3GPP 将其作为未来的主要接入手段之一。

协同中继协议模型，主要分为 AF(AmplifyandForward,放大转发)、DF(DecodeandForward,解码转发)、CF(CodedandForward,编码转发)等[3]，文献[4][5]分别讨论了存在多个中继用户时 AF 和 DF 方式的中断概率性能;文献[6]讨论了协同中继网络的容量。文献[7]对基于单中继的协同中继网络中多用户分集性能进行了研究。协同中继模型用做分布式的协同传输，可以获得较好的协同分集增益，对于不同的协同模式，文章[8]给出了协同传输时的性能增益，但大部分的研究都是基于同种网络下的中继进行研究，如分布式协同传输[9]，采用图 1 中的协同半双工模式，多个节点之间需要设计正交的空时编码信道，使得网络中各节点可以在相同的子信道内传送信息，这就需要节点间能够及时传递瞬时的信道状态信息，造成了复杂的分布式编码设计。

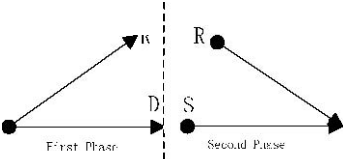


图 1 半双工分布式协同模式

1 系统模型介绍

在本文中，将中继信道模型扩展到异构网络中，网络节点间以 DF 协议进行协同传输。在中继节点上，存在多种制式的空口，因此可以接入多种不同制式的终端，而不同的制式之间可以互为 wirelessbackhaul 链路，也可以用光纤等有线技术作为传输手段。由于不同制式之间不存在干扰，且收发机独立，接收和发送信道实现频域、时域的完全正交，各制式之间就能够有效地协同工作，要达到如图 1 中的协同模式，相应的协同帧结构可以如图 2 设计：

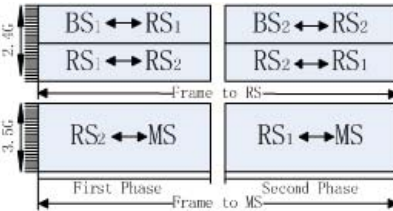


图 2 异构协同中继帧结构

如图 2 所示，协同中继节点采用免注册频段与基站和相邻中继节点通信，节点间协商好时间或频率资源，利用 3.5G 频段与终端进行同步通讯。由于中继节点能够协同两种空口制式，使得终端和基站可以采用不同制式，而在终端切换的时候，中继和中继之间通过第三方空口协同工作，可以获得更高的协同分集增益。相比传统的异构网络切换技术，如 MIH(MediaIndependentHandover)、GLL(GenericLinklayer)、UMA(UnlicensedMobileAccess)等 2.5 层技术，以及 MIP(MobileIP)、HIP(HostIdentityProtocol)、SIP(SessionInitiationProtocol)等网

络层、应用层技术，文章[10][11]做了大量研究，证明以上技术能带来有效的无线网络融合效果，但这些技术都未能给出确信的数学模型，以验证异构组网的质量，如异构切换时用户能得到的信道容量分配、中断概率等。利用协同中继原理分析异构接入的网络场景，如图 3，能够准确得出终端在切换时的信道容量、中断概率和误码率，建立可供编码和物理层设计的数学模型。

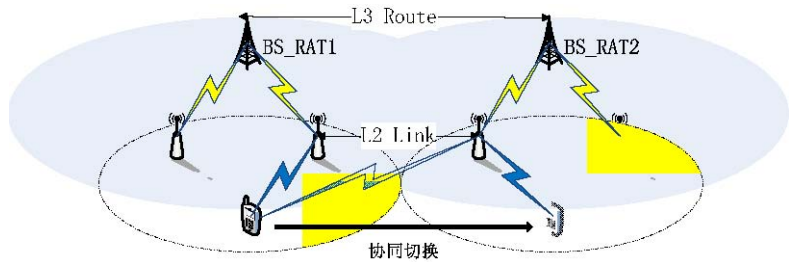


图 3 异构网络切换场景

在本文中，我们考虑这样的一个无线网络环境，其中包括多个源点（Source）、一个目的节点（Destination）、接入中继和若干辅助候选中继的集合。每个源点表示一个固定小区，相邻两个小区中，终端通过中继进行接入，由于基站与终端采用不同的无线制式，所以数据传输和切换过程，全部在中继完成；负责切换和辅助切换的两个中继站，使用同步时钟或不同频点，异步的向终端提供服务。这样既避免了基站对终端的干扰，又可以增加切换时终端所获得的信息量，从而提高异构网络环境的切换成功率。不失一般性，对网络模型的工作环境作如下假设：

- 1) 终端和基站制式不同，接入和切换过程只能由中继完成；
- 2) 终端通过 2 跳的中继进行接入，不研究复杂的中继环境；
- 3) 中继固定，且相邻中继之间存在直接链路；

目前针对协同中继模型的研究，异构性越来越受到关注，但研究主要集中在链路异构性和中继选择上，对于异构网络环境中的协同处理，尚未有系统性的模型。本文将协同中继模型推广至异构网络中，利用中继接入为结合点，研究终端在协同切换时的网络信息论模型，建立了用户容量和中断概率模型；并通过建立 DF 协同中继网络中的多用户分集模型，分析多用户场景下的网络性能，得到未来异构无线网络小区的中断概率模型和相应的中断容量；最后通过仿真验证了模型的正确性。

2 异构协同中继链路容量分析

由图 3 可知，两段链路的制式不同，可以当成完全独立正交的瑞利衰落信道，因此基站和中继站可以同时发送，终端只要能够与中继站之间准确同步，就能够利用整个空、时、频段的资源，以一种松耦合的方式工作在不同网络之间。下面分析在异构网络环境中，终端切换时，协同中继节点能够提供的性能增益。对应的网络节点模型，可由三节点模型扩展，如图 4，终端在切换时，同时接收来自两个中继的信息，由于基站和中继站位置固定，所以信道较为稳定，可以当成一个节点；中继和中继之间的信道稳定，能够提供额外的信息量；终端

接收的信息量为两个中继提供的信息量之和。可以看出信道之间完全正交，且编码复杂度未提高。

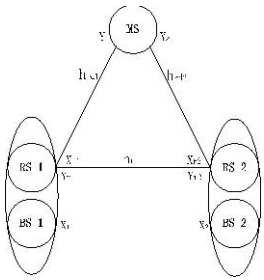


图 4 三节点扩展模型

符号间关系如下：

$$\begin{aligned} Y_{r_1}[i] &= h_{sr_1} X_1[i] + Z_{sr_1}[i]; \quad Y_1[i] = h_{rd_1} X_{r_1}[i] + Z_{rd_1}[i]; \\ Y_{r_2}[i] &= h_{sr_2} X_2[i] + Z_{sr_2}[i]; \quad Y_2[i] = h_{rd_2} X_{r_2}[i] + Z_{rd_2}[i]; \\ Y_{rr}[i] &= h_{rr} X_r[i] + Z_{rr}[i]; \end{aligned} \tag{1}$$

由[3]可知单个中继的信道容量上限可表示为：

$$C_{relay}(h_{sr}, h_{rd}) \leq \max_{P(X_1, X_r)} \min(I(X_1, X_r; Y_1), I(X_1; Y_r, Y_1 | X_r)) \tag{2}$$

其中 Y_i 表示终端从第 i 个中继节点收到的信号； Y_r 表示中继之间的交互信号； P_1, P_2, P_r 表示基站至中继，中继至终端，中继至中继的发射功率； h_{sr}, h_{rd}, h_{rr} 分别表示源节点到中继，中继到目的节点和中继节点之间的信道衰落系数，对应的路径损耗参数为 γ_{sr}, γ_{rd} 和 γ_{rr} ； Z_{sr}, Z_{rd}, Z_{rr} 分别为相应的加性白噪声。在这里，假设所有的信道衰落为平坦衰落，且为相互独立的慢衰落。由协同中继的解码转发协议可知：

$$X_r[i] = \sqrt{P_2(\overline{X(i-n)}) / P_1} \leq X_1[i-n], i = n+1, n+2, \dots, 2n \tag{3}$$

则协同中继信道的容量可表示为：

$$\begin{aligned} Y &= Y_1 + Y_2 + Y_r = h_{rd_1} X_{r_1}[i] + h_{rd_2} X_{r_2}[i] + h_{rr} X_r[i] + Z_{rd_1}[i] + Z_{rd_2}[i] + Z_{rr}[i] \\ &\leq X_1[i](h_{rd_1} + h_{rd_2} + h_{rr}) + Z_{rd_1}[i] + Z_{rd_2}[i] + Z_{rr}[i] \end{aligned} \tag{4}$$

在两个中继协同处理下的信道模型可以转化成类似图 5 的网络信息论模型，利用最大流最小割原理，可以得知该模型的理论容量为

$$C_{relay}(h_{sr}, h_{rd}, h_{rr}) = \max \min\{C_1 + C_3 + C_5, C_2 + C_3 + C_4\} \leq \frac{1}{2} \ln(1 + \frac{P_1 + P_2 + P_r}{N}) \tag{5}$$

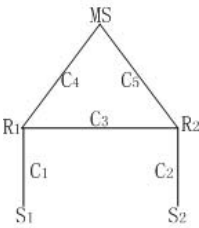


图 5 最大流-最小割模型

由公式(5)，异构协同中继的香农公式容量可以表示为：

$$C_{\text{relay}} \leq \max \min \left\{ \frac{1}{2} \ln(1 + P_1 \|h_{sr_1}\|^2 + P_r \|h_{rr}\|^2 + P_2 \|h_{rd_2}\|^2), \frac{1}{2} \ln(1 + P_1 \|h_{sr_2}\|^2 + P_r \|h_{rr}\|^2 + P_2 \|h_{rd_1}\|^2) \right\} \quad (6)$$

由于中继节点固定，则基站至中继站之间的信道衰落参数值固定，我们假设基站至两个中继节点的信道衰落参数相同，且基站的发射功率相同，中继之间的发射功率相同。则系统信噪比可表示为： $\Gamma = \Gamma + \Gamma + \Gamma$ ，协同中继信道容量可以表示为

$$I_i^{DF} = \frac{1}{2} \ln(1 + \Gamma_{sr} + \Gamma_{rr} + \Gamma_{rd_i}) \quad (7)$$

下面计算协同中继信道在异构环境中的中断概率，令 SNR 为系统的平均接收信噪比，则公式(6)中的容量计算转化为如下问题：

$$\max \Gamma_{rd} = \min(\|h_{rd_1}\|^2, \|h_{rd_2}\|^2) \overline{SNR} \quad (8)$$

求中断概率之前，需要求出容量的累积分布函数，

$$F_i^{DF}(R) = \Pr\{\|h_{sr}\|^2 + \|h_{rr}\|^2 + \min(\|h_{rd_1}\|^2, \|h_{rd_2}\|^2) < \frac{R}{SNR}\} \quad (9)$$

指数函数 $\min(\|h_{rd_1}\|^2, \|h_{rd_2}\|^2)$ 的数学期望为 $\frac{\gamma_{rd_1} \gamma_{rd_2}}{\gamma_{rd_1} + \gamma_{rd_2}}$ 其中断概率可以表示为：

$$F_i^{DF}(R) = \Pr\{I^{DF} < R\} = \Pr\{SNR^{DF} < 2^{2R} - 1\} \approx \left(\frac{2^{2R} - 1}{SNR}\right)^2 \frac{\gamma_{rd_1} + \gamma_{rd_2}}{2\gamma_{sr}\gamma_{rr}\gamma_{rd_1}\gamma_{rd_2}} \quad (10)$$

由公式(6)(10)可知，由于异构环境中，协同中继能够协同时域、频域的资源，在不增加编码复杂度的条件下，实现较为复杂的双工模式，因此信道容量有近 2 倍的提升。增加的中继之间信道，也能够有效降低中断概率。

3 多用户分集下的异构网络性能分析

由于不同制式的重叠覆盖，每个终端都处于异构网络环境中，相对于传统小区的单制式接入，新型的异构中继网络小区的性能增益可以看作小区中每个用户性能增益的极限以及多个用户的平均中断概率。在一个有多个用户的异构中继网络环境中，存在多个相互正交的用户信道时，系统中断概率为：

$$\begin{aligned} P_{out}^{DF} &= \Pr\{I^{DF} < R\} = \Pr\{SNR^{DF} < 2^{2R} - 1\} \\ &= F_i(2^{2R} - 1) = \left(\frac{2^{2R} - 1}{SNR}\right)^{2N} \prod_{i=1, j=i+1}^N \frac{\gamma_{rd_i} + \gamma_{rd_j}}{2\gamma_{sr}\gamma_{rr}\gamma_{rd_i}\gamma_{rd_j}} \end{aligned} \quad (11)$$

利用系统的中断概率，使用以下公式，可以得出异构网络能够提供的网络容量：

$$\begin{aligned} P_{out} &= \varepsilon \\ F(x) &= \Pr\{\|h\|^2 > x\} \\ C_{out} &= \ln(1 + F^{-1}(1 - \varepsilon) \overline{SNR}) \end{aligned} \quad (12)$$

由公式可知，如用户之间信道完全独立正交，系统性能随着用户数的增加，中断概率无限逼近零，分集阶数达到 $2N$ 。由于协同中继的存在，使得异构网络之间可以互为补充，终端得以在不同制式的网络中漫游，且能获得较高的容量增益和较低的中断概率。下面通过仿真详细比较理论分析结果。

4 数值分析

系统仿真环境基于同济大学宽带无线通信与多媒体研究室的下一代宽带无线通信网络的仿真验证平台，用以验证协同中继节点在同构和异构网络中的性能差异，以及分析多用户分集对异构网络环境产生的性能增益。无线网络环境为两个蜂窝小区、两个协同中继节点、10个终端用户，小区半径设置为 1000m，为了简化中继选择过程，终端切换时直接选择相邻中继，信道环境为瑞利慢衰落信道。

图 6 给出了异构协同中继模型的香农公式容量，图中两条曲线分别表示协同中继信道在同构环境和异构环境中取得的香农容量上限。由于切换时多中继协作且各 DF 信道之间严格独立正交，因此，因此在低信噪比区域，理论上的异构中继信道容量相比同构中环境下的有 2bit/s/hz 的增长，且随着信噪比的增长，协同中继在异构网络环境中的性能曲线斜率更小。

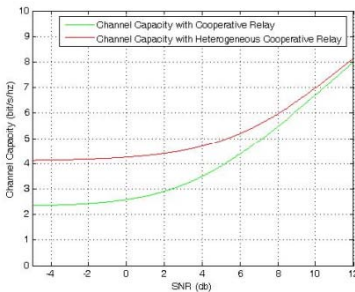


图 6 同构和异构环境中的中继信道容量

图 7 和图 8 分别给出了基于协同中继模型的中断概率和中断容量，以及给出了多用户分集对两种指标的影响。图 7 中协同中继在同构环境和异构环境产生的中断概率在同一个数量级内，由公式(11)可知，多用户分集对中断容量的影响有极限，因此在这里用户数目分别选择 3、5、10，能产生较为明显的差别，以下系统中断容量中的用户数也是同理选择。随着用户数目的增加，中断概率分布在 10^{-5} — 10^{-20} 之间变化。图 8 中的各线表示协同中继在同构和异构环境中所能达到的中断容量，以及多用户分集对异构系统中断容量的影响。由图 8 可知，在低信噪比区域，异构协同中继性能比同构环境下提升 1db 左右，多用户分集也能带来 0.5db 的性能增益。

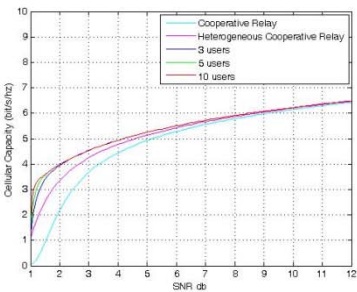
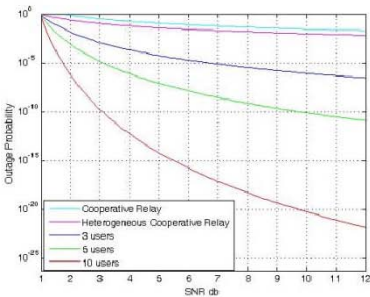


图 7 中断概率比较及多用户分集性能 图 8 中断容量比较及多用户分集性能

5 结论

本文给出了异构中继接入的系统参考模型, 对未来网络的规划部署做了初步探讨, 验证了异构网络环境下, 该模型的性能增益和多用户分集对系统模型的影响。这种模型不仅可以实现较好的切换性能, 还可以和 MIP、SIP 等应用协议研究紧密结合, 实现应用业务的无缝切换。未来的研究方向将围绕小区干扰建模分析和切换的 QoS 性能分析, 如时延、吞吐量、误码率等。

参 考 文 献

- [1] Beaubrun Ronald, Holland Oliver, Aghvami Hamid. Heterogeneous Wireless Access Networks Architectures and Protocols[M]. Canada. Springer Science 2008, ISBN: 978-0-387-09776-3.
- [2] Vodafone, Ericsson, Alcatel-Lucent. Type II relay frame-work definition[EB/OL]. 3GPP, R1-091632, March 2009.
- [3] Laneman Nicholas, Tse David, Wornell Gregory. Cooperative Diversity in Wireless Networks: Efficient Protocols and Outage Behavior[J]. IEEE Trans Information Theory, VOL. 50, NO. 12, DECEMBER 2004.
- [4] Laneman J N , Wornell GW. Distributed space-time-coded protocols for exploiting cooperative diversity in wireless networks [J] . IEEE Trans Information Theory , 2003 , 49 (10) :241422425.
- [5] Zhao Y, Adve R , Lim T J . Improving amplify-and-forward relay networks : optimal power allocation versus selection [J] . IEEE Trans Wireless Commu , 2007 , 6 (8) :311423123.
- [6] Zhu Song , Wang Wenbo , Cheng Yu. Capacity analysis for heterogeneous and homogeneous cooperative relay channels[J] . Journal of Beijing University of Posts and Telecommunications , 2008 , 31 (4) : 77-81.
- [7] Zhang Xing , Chen Shuping , Wang Wenbo. Performance analysis of multiuser diversity in multiuser two-hop cooperative relay wireless networks[C] . IEEE ICC. Beijing : IEEE Press , 2008 : 337-341.
- [8] Fizek Frank, Katz Marcos. COOPERATION IN WIRELESS NETWORKS: PRINCIPLES AND APPLICATIONS[M]. Netherlands. Springer, 2006.
- [9] Hunter Todd Edward. CODED COOPERATION: A NEW FRAMEWORK FOR USER. COOPERATION IN WIRELESS NETWORKS[D]. Univ. Texas at Dallas, Richardson, 2004.
- [10] Jae-Han Seol; Jong-Moon Chung. IEEE 802.21 MIH based Handover for Next Generation Mobile Communication Systems[C]. Innovations in Information Technology, 2007. Innovations '07. 4th International Conference. 18-20 Nov. 2007 Page(s):431 – 435.
- [11] Banerjee Wei Wu. SIP-based vertical handoff between WWAN and WIAN[J]. IEEE Wireless Commun, Special Issue: Toward Seamless Internetworking of Wireless LAN and Cellular Networks, 2005, 12(3):66-72.

作者简介:

高智伟, 1981 年出生, 男, 江西鹰潭人, 同济大学控制理论与控制工程博士生, 主要研究方向为协同原理在异构网络环境中的应用。作者简介放基金项目下方。按以下顺序给出全部作者的简介: 姓名(出生年-), 性别, 民族(汉族可省略), 籍贯(省、市/县), 学位, 单位, 职称, 职务, 简历及主要研究方向。

联系地址: 上海市曹安公路 4800 号同济大学 13 号楼 223 室; 邮编: 201804; 手机: +86-13501935705; 电子邮箱: golgo@live.cn

分簇无线传感器网络中最小退避窗口自适应机制

林颖 许力 林晖

(福建师范大学数学与计算机科学学院, 福建, 福州, 350007)

摘要: 为了向无线传感器网络节点提供区分服务, 文章提出了一种支持区分服务并且具有一定公平性的媒体接入控制机制-ABW-DS。在分簇的传感器网络中, 首先由簇头计算不同优先级节点的吞吐量, 然后设定一定的阈值, 在不把低优先级节点的带宽全部占用的前提下, 高优先级节点能获得较高的吞吐量的保证。

关键词: 区分服务; 无线传感器网络; 公平性; 最小退避窗口

An Adaptive Backoff Windows Scheme in Cluster-based Wireless Sensor Networks

Lin Ying Xu Li Lin Hui

(Fujian Normal University, School of Mathematics and Computer Science, Fuzhou, 350007)

Abstract: For providing differentiated service in Cluster-based wireless sensor networks, an adaptive backoff windows scheme-ABW-DS which provides the differentiated service and at the same time provides the fairness is proposed in this paper. In cluster-based wireless sensor networks, the cluster head calculate two different throughputs of the nodes with service classes, through enacting a threshold, the proposed protocol can control that the nodes with high priority can win some bandwidth but can't embezzle the bandwidth of the nodes with low priority.

Keywords: Differentiated Service; Wireless sensor networks; Fairness; backoff windows

1 介绍

微传感器技术、微电子技术、无线通信技术以及计算技术的进步, 极大地推动了集信息采集、处理、无线传输等功能于一体的无线传感器网络(wireless sensor networks, 简称 WSNs)的发展。WSNs 正在给人类生活和生产的各个领域带来深远影响。在国防军事、医疗卫生、环境监测、城市交通以及空间探索等领域具有广阔的应用前景^[1]。

在 WSNs 中, 由于多跳和临时性的特点, 一般不存在一个中心协调装置, 而多采用异步的、分布式的链路层接入方式, 主要是 802.11MAC 接入方式。目前, 基于 802.11 分布式控制的接入方式, 没有对节点级别以及数据业务的级别加以区分, 存在公平性差以及不能很好支持区分服务的缺点。

Y. Xiao^[2]通过仿真实验证明, 当网络通过设置不同的最小退避窗口值来区分节点的优先级时, 即高优先级节点设置较小的最小退避窗口值, 低优先级节点设置较大的最小退避窗口值,

随着低优先级节点最小退避窗口的增加，高优先级节点可以“盗用”低优先级节点的带宽，出现了不公平现象，如图 1 所示，并且高优先级节点的饱和延迟减少，而低优先级节点的饱和延迟急剧增加，如图 2 所示。随着低优先级节点继续扩大最小退避窗口值，有可能会出现低优先级节点根本没有带宽来传输数据的情况。这时，就出现了高优先级节点过度“盗用”低优先级节点带宽的情况，即出现了极度不公平的情况。多数研究者研究了如何保持节点之间能公平的获得相同的带宽，如 MACAW^[3]，GPS^[4]，NEMA-MAC^[5]等协议，并未同时考虑节点分区服务的情况。

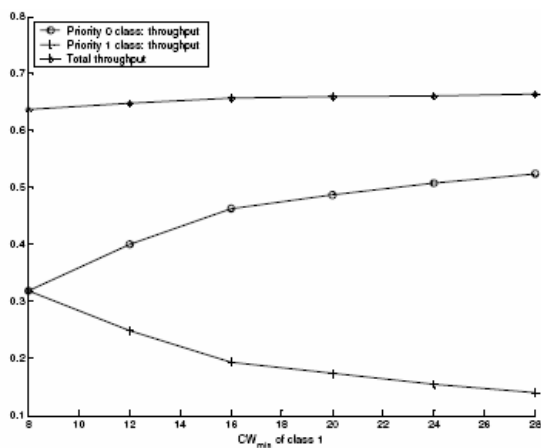


图 1 随着低优先级节点最小窗口变化时的饱和吞吐量 (Mbit/s)

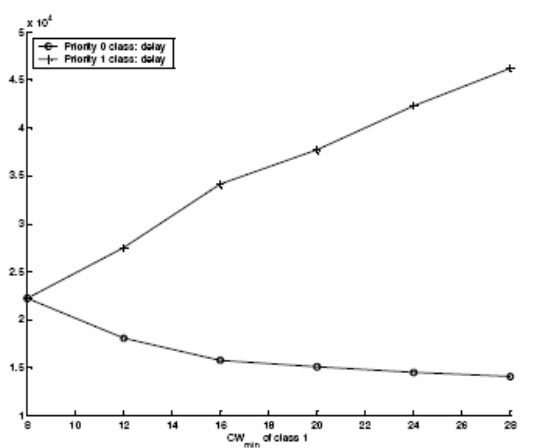


图 2 随着低优先级节点最小退避窗口变化时的饱和延迟 (μs)

本文兼顾公平性与区分服务，从平衡公平性与服务区分出发，通过改进 802.11MAC 层协议，在网络已分簇的情况下，由簇头节点存储簇内不同优先级节点的吞吐量信息来改变不同级别节点的退避窗口值的大小，从而实现保证高优先级节点占用一定的带宽的情况下，不会完全地“盗用”低优先级节点的带宽。

2 基于区分服务和带宽保证的最小退避窗口自适应机制-ABW-DS

假设网络中一共有 $2n$ 个节点，高优先级节点与低优先级节点分别为 n 个。在已分簇的网络中，簇头维持分簇结构。每个簇头节点负责簇成员节点的通信，簇内节点要想跟 SINK 节点通信，必须通过簇头节点才能到达 SINK 节点。如图 3 所示。高优先级节点的最小退避窗口的当前值为 CW_0 ，初始值为 W_0 ；低优先级节点的最小退避窗口的当前值为 CW_1 ，初始值为 W_1 ，并且 $W_0 < W_1$ 。

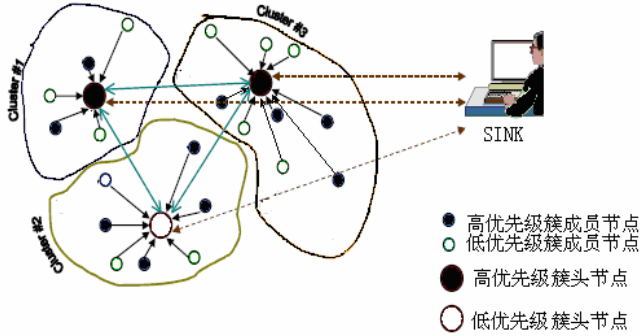


图 3 分簇无线传感器网络

设一个簇内，高优先级节点个数为 m_0 ，低优先级节点个数为 m_1 。簇头节点 j 在本地保存两个本地吞吐量信息 $TP_{L_j}^i$ ($i=0, 1$)，其中， $TP_{L_j}^0$ 为 m_0 个高优先级节点的总本地吞吐量， $TP_{L_j}^1$ 为 m_1 个低优先级节点的总本地吞吐量。记 $TP_{C_j}^0$ 是簇内的高优先级节点的当前总吞吐量， $TP_{C_j}^1$ 为簇内的低优先级节点的当前总吞吐量。簇头节点 j 以时间 T 为周期，分别统计 $TP_{C_j}^0$ 和 $TP_{C_j}^1$ 。同时，簇头节点保存的本地吞吐量的信息为当前吞吐量和先前吞吐量的加权和，簇头节点采用公式 (1) 的迭代方式来计算两个本地吞吐量的值：

$$TP_{L_j}^i \leftarrow \alpha \cdot TP_{L_j}^i + \beta \cdot TP_{C_j}^i \quad i=0,1 \quad (1)$$

其中， α 和 β 为常数，表示了历史信息的重要性， $TP_{L_j}^i$ ($i=0, 1$) 的初始值为 0。

簇头节点计算出 $TP_{L_j}^i$ ($i=0, 1$) 后，根据以下的 3 种情况判断高优先级节点和低优先级节点的当前最小退避窗口的变化。

情况 1: 当 $\frac{TP_{L_j}^1}{m_1} : \frac{TP_{L_j}^0}{m_0} < \gamma_0$ 时，其中 γ_0 是一个不大于 1 的常数，则在该簇内，高优先级节点已经占用了大部分的带宽，低优先级节点带宽很小，这种情况说明高优先级节点过度“盗用”低优先级节点的带宽，即最不公平的情况。此时，高优先级节点采用公式 (2) 来增加它的最小退避窗口来发送数据，使得低优先级的节点有更多的机会来接入信道。

$$CW_0 \leftarrow \lfloor CW_0 \cdot \text{random}[\alpha_1, \alpha_2] \rfloor \quad (2)$$

而低优先级节点就必须采用小的退避窗口来竞争信道，如公式 (3)。

$$CW_1 \leftarrow \lfloor CW_1 \cdot \text{random}[\beta_1, \beta_2] \rfloor \quad (3)$$

其中, β_1 、 β_2 、 α_1 、 α_2 为正数, $\beta_1 < \beta_2 \leq \alpha_1 < \alpha_2$ 。记号 $\lfloor x \rfloor$ 表示不大于 x 的最大整数。记号 $\text{random}[x, y]$ 表示一个均匀分布于区间 $[x, y]$ 上的随机数,

情况 2: 当 $\gamma_1 > \frac{TP_{L_j}^1}{m_1} : \frac{TP_{L_j}^0}{m_0} > \gamma_0$ 时, 其中 γ_1 是一个大于等于 1 的常数, 说明在该簇内高优先级节点仅占用了小部分低优先级节点的带宽, 属于节点区分服务级别时的正常现象, 则高、低优先级节点都不需要改变其本身的最小退避窗口值。

情况 3: 当 $\gamma_1 < \frac{TP_{L_j}^1}{m_1} : \frac{TP_{L_j}^0}{m_0}$ 时, 说明在该簇内, 高优先级节点没有占用大部分的带宽, 反而因为最小退避窗口的调整而损失了自身应该占用的带宽。此种情况也是不应当出现的情况, 那么高优先级节点和低优先级节点都把最小退避窗口恢复成初始值。

当簇头节点判断簇成员的最小退避窗口应该如何变化后, 它以 T 时间为周期, 发送广播信息。广播帧的格式如图 4 所示, 其中 B 代表该信息为广播信息, 占用了一个字节, 后面跟着 4 个比特的控制位。4 个比特中, 前 2 个比特表示高优先级节点的最小退避窗口变化情况, 后 2 个比特表示低优先级节点的最小退避窗口变化情况。其中, 00 表示最小退避窗口值不变, 01 表示最小退避窗口需要按照情况 1 来变, 10 表示最小退避窗口要按情况 3 来变化。

B	0/1	0/1	0/1	0/1
---	-----	-----	-----	-----

图 4 簇头的广播帧的格式

当簇成员节点收到广播帧后, 根据 β_1 、 β_2 、 α_1 、 α_2 值来改变自身的最小退避窗口。

3 仿真分析

本文通过仿真实验来对比在基于 802.11DCF 的协议和基于本文中提出的 ABW-DS 协议的网络吞吐量分析。物理层采用 802.11 协议的参数, 数据速率取 6Mbit/s, 控制信息的速率也为 6Mbit/s, 帧长度固定为 1024 字节。高优先级节点的最小退避窗口为 8, 低优先级节点的最小退避窗口值从 8 变化到 18。 n 取 10, $\beta_1 = 1$, $\beta_2 = 1.5$, $\alpha_1 = 1.5$, $\alpha_2 = 2$, $\gamma_0 = 0.5$, $\gamma_1 = 1$, $T = 0.5s$ 。

从图 5 中我们可以看出, 当采用 802.11 协议时, 随着低级别节点的最小退避窗口的增加, 低优先级节点的吞吐量被高优先级节点“盗用”。特别是当低优先级节点的窗口为 13 时, 高优先级节点已经占用了整个带宽的 65%。而采用了 ABW-DS 协议, 我们发现, 当高优先级节点占用的带宽未超过阈值 γ_0 , 则与 802.11DCF 是一致的, 一旦出现了高优先级节点带宽超过阈值 γ_0 的情况, 即低优先级节点的最小退避窗口值为 13 时, ABW-DS 协议可以使得高优先级节点停止进一步“盗用”低优先级节点的带宽, 但是却不影响高优先级节点获得比低优先级节点更多的带宽, 同时保证了低优先级节点有带宽可以发送数据。

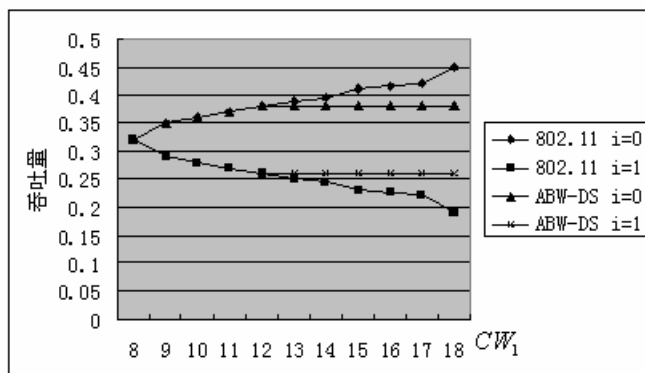


图 5 随着 CW_1 变化的饱和吞吐量

4 小结

本文在分簇无线传感器网络中，簇头计算不同优先级节点的吞吐量，并且设定一个阈值，使得网络中的高优先级节点在获得相当大部分的带宽的同时，保证了低优先级节点能够有一定的带宽传输数据，实现公平性与区分服务的平衡。仿真实验证明，ABW-DS 确实能使得低优先级节点在最小退避窗口持续扩大的同时，还能获得一定的带宽，而高优先级节点虽然不能完全“盗用”低优先级节点的带宽，但是也获得了相当比例的带宽，实现整个网络在保持区分服务的情况下，保证所有的节点都有带宽可以传输数据。

参 考 文 献

- [1] Callaway EH. Wireless Sensor Networks: Architectures and Protocols[J]. San Diego: CRC Press, 2003:1-17
- [2] Y.Xiao.Performance analysis of IEEE 802.11e EDCF under saturation condition[J]. Proc. of IEEE ICC 2004 1(20-24): 170-174.
- [3] V. Bharghavan, A. Demers, S. Shenker, and L. Zhang. MACAW: A media access protocol for wireless LANs[J]. August 1994 :34-46.
- [4] N. H. Vaidya and P. Bahl. Fair scheduling in broadcast environments[J]. *Tech. Rep. MSR-TR-99-61, Microsoft Research*
- [5] Yongkang Xiao, Lin Zhang, Xiuming Shan, Yong Ren, and Zhengxin Ma.Neighbor-medium-aware MAC protocol with fairness guarantee for wireless ad hoc networks[J], IEICE Transactions on Communication , 2004, 87(9): 2738-2746.

一种解决认知无线电频谱分配的量子遗传算法

郭 明 李 飞

(南京邮电大学通信与信息工程学院, 南京 21003)

摘 要: 提出了一种量子遗传算法的认知无线电频谱分配算法, 通过仿真比较了本文算法与经典的遗传算法的性能。结果表明基于这种量子遗传算法的频谱分配算法性能明显优于经典的遗传算法, 它能更好地实现网络效益最大化; 当用户数和频带数较少时, 量子遗传算法在找到最优解的前提下可以使进化代数很少, 而经典的遗传算法所得到的解与理想最优解有所偏差, 而且偏差值较大。

关键词: 遗传算法; 量子遗传算法; 认知无线电; 频谱分配模型

A Solution About Quantum Genetic Algorithm to the Allocation of Radio Spectrum Cognitive

Guo Ming Li Fei

(School of Telecommunication and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing 210003, China)

Abstract: The paper present a quantum genetic algorithm cognitive radio spectrum allocation algorithm, through simulations compare the proposed algorithm with the classical genetic algorithm performance. The results show that genetic algorithm based on this quantum of spectrum allocation algorithm is superior to the classical genetic algorithm, it can better achieve network efficiency maximization; when the number of users and fewer bands, the quantum genetic algorithm to find the optimal solution algebra under the premise of evolution can make small, while the classical genetic algorithm and the ideal solution obtained deviate from the optimal solution, but a larger deviation.

Keywords: Genetic Algorithm; Quantum Genetic Algorithm; Cognitive Radio; Spectrum Allocation Model

1 引言

近年来, 随着无线通信技术的飞速发展, 尤其是无线局域网(WLAN)技术、无线个人域网络(WPAN)技术的发展, 越来越多的人通过这些技术以无线的方式接入互联网, 使得无线频谱成为越来越紧缺的资源。为了解决频谱资源匮乏的问题, 基本思想就是尽量提高现有的频谱利用率, 于是人们提出了这样的一种方法: 在已授权频段没有或只有很少的通信业务在活动的情况下, 具有认知功能的无线通信设备可以侦测出在空域、时域和频域中出现的可被利用的频谱资源, 并按照某种“伺机(Opportunistic Way)”方式工作在这些已授权的频段内, 合理

的使用这些“频谱空洞”。这种使无线通信设备具有发现“频谱空洞”并合理利用它们的能力构成了认知无线电(CognitiveRadio, CR)的核心思想^[1]。

认知无线电智能感应外部无线环境，以达到高可靠通信和使频谱利用率最大^[2]。本文着眼于研究空闲频谱资源在次用户间的分配问题。到目前为止，动态频谱分配方法主要包括博弈论、拍卖理论、议价机制、图论着色等。文献[3][4]提出了一种认知无线电频谱分配模型，本文在讨论这种频谱分配模型基础上，提出了一种基于量子遗传算法(quantum genetic algorithm, 简称 QGA) 的频谱分配方法，通过仿真对基于 QGA 的频谱分配算法和基于经典的遗传算法性能进行了比较，仿真结果验证了本文所提算法的高效性和优越性。

2 量子遗传算法

量子遗传算法^{【5-8】} (QGA—Quantum Genetic Algorithm) 是将经典遗传算法 (GA—Genetic Algorithm) 与量子计算理论相结合而发展起来的一种新的遗传算法。

QGA 使用一种基于量子比特的编码方式，即用一对复数定义一个量子比特位，一个有 k 个量子比特位的系统描述为

$$\begin{bmatrix} \alpha_1|\alpha_2|\cdots|\alpha_k \\ \beta_1|\beta_2|\cdots|\beta_k \end{bmatrix}, \qquad |\alpha_i|^2+|\beta_i|^2=1, \qquad i=1,2,\cdots,k$$

其中 α_i, β_i 是两个复数，是第 i 个量子比特的概率幅，其模满足归一化条件， $|\alpha_i|^2$ 表示测量时发现 $|0\rangle$ 的概率， $|\beta_i|^2$ 表示发现 $|1\rangle$ 的概率。这种表示可表征任意的线性叠加态，设一个染色体包含 k 位量子位，则其编码形式为：

$$\mathbf{q}=\begin{bmatrix} \alpha_1|\alpha_2|\cdots|\alpha_k \\ \beta_1|\beta_2|\cdots|\beta_k \end{bmatrix}$$

这样编码后，QGA 中染色体种群可以表示为 $\mathbf{Q}(g)=\{\mathbf{q}_1^g, \cdots, \mathbf{q}_p^g\}$, P 为种群大小， g 是进化代数， $\mathbf{q}_i^g=\begin{bmatrix} \alpha_{i1}^g & \cdots & \alpha_{ik}^g \\ \beta_{i1}^g & \cdots & \beta_{ik}^g \end{bmatrix}$ 。

测量将改变量子位的状态，使其从 $|0\rangle$ 和 $|1\rangle$ 的叠加态塌缩到与观测结果相应的特定状态. 对种群 $\mathbf{Q}(g)$ 中的各个个体实施一次测量，将得到一组状态 $\mathbf{P}(g)$. $\mathbf{P}(g)$ 是一组二进制解，其中每一个二进制解有长度为 k 的二进制串组成。，并且 \mathbf{p}_i^g 中第 j 位的取值通过 \mathbf{q}_i^g 中第 j 位的量子位的概率确定。即随机产生一个 0 到 1 之间的数，与其比较大小，小的话取 0，反之取 1. 种群一开始均初始化为 $1/\sqrt{2}$, 表示在初始搜索时所有状态均以相同的概率进行叠加. 文中使用的量子旋转门的更新操作如下所示：

$$\begin{pmatrix} \alpha'_{ij} \\ \beta'_{ij} \end{pmatrix} = \begin{pmatrix} \cos\theta_{ij} & -\sin\theta_{ij} \\ \sin\theta_{ij} & \cos\theta_{ij} \end{pmatrix} \begin{pmatrix} \alpha_{ij} \\ \beta_{ij} \end{pmatrix}$$

其中， $\begin{pmatrix} \alpha_{ij} \\ \beta_{ij} \end{pmatrix}$ 是染色体 i 第 j 量子位， $\begin{pmatrix} \alpha'_{ij} \\ \beta'_{ij} \end{pmatrix}$ 是更新后的结果。 θ_{ij} 是量子旋转门的旋转角，

本文采用的公式是 $\theta_{ij}=\Delta \theta_{ij} \cdot \mathbf{S}(\alpha_{ij}, \beta_{ij})$ ， $\Delta \theta_{ij}$ 和 $\mathbf{S}(\alpha_{ij}, \beta_{ij})$ 分别是旋转的角度和方

向, 具体的步骤见文献。该调整策略将待调整个体当前的测量值的适应度, 使得概率幅对 $\begin{pmatrix} \alpha_{ij} \\ \beta_{ij} \end{pmatrix}$ 向着有利于 p_{ij} 出现的方向演化 (当适应度值大于已保存的最优解的适应度值); 反之, 向着反方向演化^[7]。

QGA 基本流程图如图 1 所示算法终止条件一般以是否达到最大进化代数 G 作为衡量标准。随着进化的进行, 种群的解逐渐向最优解收敛。

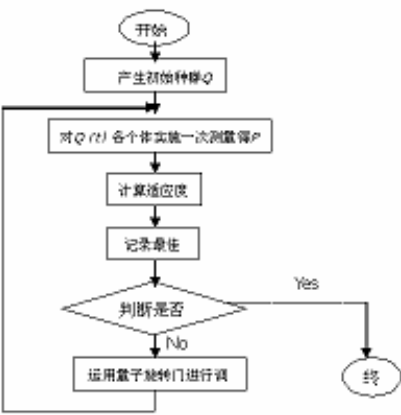


图 1 QGA 流程图

3 认知无线电频谱分配模型

认知无线电频谱分配模型^[4]由可用频谱矩阵、效益矩阵、干扰矩阵和无干扰分配矩阵描述。假定 N 为次用户数, 各个次用户分别用标号 $0, 1, \dots, N-1$ 表示, M 为认知无线电可用频带数, 频带间相互正交, 各个频带分别用 $0, 1, \dots, M-1$ 表示。

可用频谱矩阵是指在某个空间、某个时间主用户不占用的频谱。频谱对于次用户是否空闲可用频谱矩阵 L 表示: $L = \{l_{n,m}\}_{N \times M}$ 。 $l_{n,m} = 1$ 表示用户 n 可以使用频带 m , $l_{n,m} = 0$ 表示用户 n 不能使用频带 m 。

不同的次用户采用的发射功率、空时特性、调制技术等有所不同, 在不同的频谱上具有不同的传输效益!。用户获得的效益用效益矩阵 B 表示: $B = \{b_{m \times n}\}_{N \times M}$, $b_{n \times m}$ 表示用户 n 在频带 m 上获得的效益, 如频谱利用率、最大流量等。

若有不同认知无线电发射端用户在同时使用同一频带时, 两用户之间会产生干扰, 则它们是不可以同时使用频带的。在检测是否会产生干扰时。主要是从认知用户的保护半径考虑的, 用户之间的干扰用干扰矩阵 C 表示 $C = \{C_{n,k,m}\}$ 。 $C_{n,k,m} = 1$ 表示用户 n 和用户 k 同时使用频带 m 时会产生干扰, $C_{n,k,m} = 0$ 则表示不会产生干扰。当 $n=k$ 时, $C_{n,n,k} = 1 - l_{n,m}$ 实际应用中, 由于认知无线电系统进行频谱分配的时间相对于频谱环境变化的时间很短, 因此假定用户地理位置、可用频谱资源等是静态的, 即矩阵 LBC 在一个分配周期内维持不变。

另外还有无干扰分配矩阵 $A = \{a_{n,m}\}$, $a_{n,m} = 1$ 表示频带 m 已经分配给用户 n , 等于 0 就是没有被分配。

由上面几个矩阵得出，用户获得的总效益就可以用效益向量表示为：
 $R = \{\beta_n = \sum_{m=0}^{M-1} a_{n*m} b_{n*m}\}_{N*1}$ ，其中的 β_n 为用户 n 获得的总效益，认知无线电频谱分配目标即最大化网络效益 $U(R)$ 本文里面的 $U(R) = \sum_{n=0}^{N-1} \beta_n$ 就是最大化整个网络效益的总和，另外在程序的调试中还把因为用户数的增多对信道的损耗因素考虑进去，从而对网络效益做了更为接近事实的模拟。

4 QGA运用于频谱分配

本文提出的基于 QGA 的频谱分配算法中，每一条染色体测量后的二进制串表示一种可能的频谱分配。由于与可用频谱矩阵 L 中值为 0 的元素位置相对应的无干扰频谱分配矩阵 A 中的元素值必定为 0，若将 A 中所有元素均用量子位表示，将使染色体中包含大量冗余，所以本文仅对与 L 中值为 1 的元素位置对应的 A 中的元素进行量子位编码，故染色体中量子位个数等于 L 中值为 1 的元素个数。

衡量染色体测量值性能的适应度函数与目标函数相对应，由于频谱分配所要实现的目标是最大化网络效益 $U(R)$ ，故本文直接将 $U(R)$ 作为适应度函数！。随着进化的进行，种群中个体测量值适应度不断增加，达到最大迭代次数时，算法终止，此时将 $B(g)$ 保存的解映射为 A 的形式，即得到了最佳的频谱分配。

由上述得到的算法流程如下：

利用给定的矩阵 LBC 求出种群规模 P 和染色体的量子位数 l ，然后记录 L 中值为 1 的 n 和 m ，并按照 n 递增， m 递增的顺序排列初始化种群。

对 $Q(g)$ 测量得到一种状态 $P(g)$ 。

讲 $P(g)$ 映射到矩阵 A 中去。

对 $P(g)$ 进行适应度评价，并把最优解保存到 $B(g)$ 。

如果到了进化代数则终止，得到最优解。否则继续

继续下一代 $g=g+1$ 。

重复 3 4 5 步

使用量子旋转门对种群进行更新，得到 $Q(g)$ 转到 6。

5 性能仿真和结果分析

本文的 QGA 参数取值为 $P=20$ ，最大进化代数是 60。图 2 给出了 $N=8$ ， $M=4$ 的时候不同初始条件下的结果，不同实验中 LBC 不同，同一次实验中两个算法所采用的 LBC 相同， BLC 根据文献[4]附录 I 提供的伪码仿真产生。从图 2 中可以看出，60 次实验中基于 QGA 的频谱分配算法所得到的网络效益多数明显大于经典遗传算法获得的网络效益，仅有少数点较为接近，说明了基于 QGA 的频谱分配算法的有效性和优越性。

图 3 给出了 $M=4$ 时平均效益随次用户数 N 的变化曲线，平均效益等于网络效益除以 N 。由图 2 可知，随着 N 的增加，平均效益呈递减趋势。基于 QGA 的频谱分配算法所得到的平均效益均大于基于 GA 的频谱分配算法。

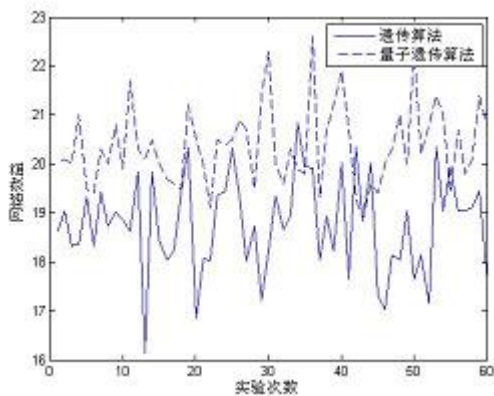


图 2

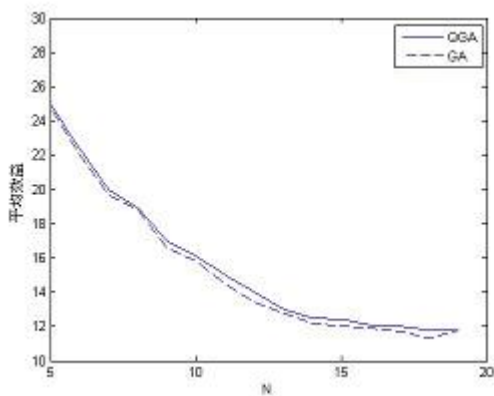


图 3

6 结论

如何合理有效地分配次用户所需的频谱是实现动态频谱接入的关键技术。认知无线电频谱分配模型可以表示为一个优化问题。本文使用 QGA 求解该优化问题，提出了基于 QGA 的认知无线电频谱分配方法，并与经典的遗传算法进行了性能比较。仿真结果表明基于 QGA 的算法能更好地实现网络效益的最大化，较小进化次数下就能找到理想最优解，验证了本文所提方法的有效性。

参 考 文 献

- [1] HayKin S 2005 IEEE J.Sel Area.Comm.23201
- [2] Zhao J Z, Zheng S L,Shang J N,Kong X Z 2007 Acta Phys.Sin.566760
- [3] Zheng H,Peng C 2005 Proc.40th Annual IEEE International Conference On Communication
- [4] Peng C,Zheng H,Zhao B Y 2006 ACM Mobile Networks And Application
- [5] Narayanna A., Moore M., "Quantum-inspired Genetic Algorithms", Proceedings of IEEE International Conference on Evolutional Evolution, 61-66, 1996

- [6] Han K.-H., Kim J.-H., “Genetic Quantum Algorithm and its Application to Combinatorial Optimization”, Proceedings of the 2000 IEEE Congress on Evolutionary Computation, 1354-1360, 2000
- [7] Han K.-H., Park K.-H., Lee C.-H., Kim J.-H., “Parallel Quantum-inspired Genetic Algorithm for combinatorial Optimization Problem”, Proceedings of the 2001 IEEE Congress on Evolutionary Computation, 1422-1429, 2001
- [8] Han K.-H., Kim J.-H., “Quantum-Inspired Evolutionary Algorithm for a Class of Combinatorial Optimization”, IEEE Trans. on Evolutionary Computation, Vol. 6(6): 580-593, 2002

一种求解最短路径路由问题的量子遗传算法

刘 欣 李 飞 张 杰

(南京邮电大学通信与信息工程学院, 南京 21003)

摘 要: 论述了用量子遗传算法解决计算机路由选择中的最短路径路由问题, 采用量子比特编码, 利用量子旋转门操作实现量子染色体的更新操作。计算机仿真实验表明, 与传统的遗传算法比较, 该算法快速有效、可靠性高。

关键词: 遗传算法; 量子遗传算法; 路由选择; 最短路径路由

A Quantum Genetic Algorithm for Shortest Path Routing Problem

Liu Xin Li Fei Zhang Jie

(School of Telecommunication and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing 210003, China)

Abstract: This paper presents a quantum genetic algorithm approach to the shortest path (SP) routing problem. Quantum bit (qubit) has been used for encoding the problem. The quantum chromosomes are updated by the quantum gate operation. Computer simulations show that the Quantum Genetic Algorithm (QGA) exhibits a better quality of solution (route optimality) than the conventional Genetic Algorithm (GA).

Keywords: Genetic algorithm; Quantum genetic algorithm; Routing; Shortest path routing

近年来, 随着 Internet 网络和移动 Ad-hoc 网络的飞速发展, 路由选择的研究已成为网络领域中的一类重要课题。一些传统的最短路径 (Shortest Path, SP) 搜索算法: 宽度优先搜索算法、Dijkstra 算法、Bellman-Ford 算法等, 它们在有基础设施的无线网络和有线网络方面在多项式时间内能较好地解决最短路径问题, 但在实时通信环境、高动态的拓扑结构以及 QoS 要求的网络结构下, 通常需要同时寻找出一组最短或次最短路径作为方案评价和选择的依据^[1-2]。

经典遗传算法 (GA—Genetic Algorithm)^[3] 是一类借鉴生物界自然选择和物种遗传机制而形成的并行随机群体搜索算法。GA 中常用的遗传操作有交叉、变异和选择等。量子遗传算法 (QGA—Quantum Genetic Algorithm)^[4] 是将 GA 与量子计算理论相结合而发展起来的一种新的遗传算法。文献[4]中提出的 QGA 以量子计算的一些概念和理论为基础, 用量子比特编码来表示染色体, 用量子门作用和量子门更新来完成进化搜索, 使得该算法将来在量子计算机上执行成为可能。QGA 具有种群规模小而不影响算法性能、同时兼有开发和探索的能力、收敛速度快等特点, 用于求解组合优化问题时表现出明显优于经典 GA 的优势。

对于用遗传算法解决 SP 路由问题^[5,6]、多播路由问题、动态路由问题等都属于组合优化问

题。文献[5]采用可变长度的染色体的编码方式。文献[6]采用基于优先权的编码方式。

1 网络模型

在研究路由问题时，一个网络可表示成一个加权图 $G(V,E)$ ，其中 V 表示节点集， E 表示连接节点的通信链路集。即无向图 G 由一组点 $V=\{1, 2, \dots, n\}$ 和一组连接 V 中节点的边 $E \in V \times V$ 组成。 C_{ij} 为链路 (i,j) 的代价，代价矩阵为 $C=[C_{ij}]$ 。从节点 i 到 j 的路径是 E 中的一个边序列 $(i, l), (l, m), \dots, (k, j)$ 。路径中没有有一个节点出现的次数多于 2。路径也可以用一个节点序列 (i, l, m, \dots, k, j) 等价表示。

令 1 表示路径的初始节点， n 表示路径的终节点， x_{ij} 是如下定义的指示变量：

$$x_{ij} = \begin{cases} 1 & \text{如果边}(i,j)\text{在路径上} \\ 0 & \text{其他} \end{cases}$$

最短路径问题可以表示为：

$$\min \sum_i \sum_j c_{ij} x_{ij} \quad (1)$$

$$\text{s.t.} \quad \sum_j x_{ij} \leq 2, \quad \forall i \in V \quad (2)$$

$$\sum_{j \neq k} x_{ij} \geq x_{ik}, \quad \forall (i,k) \in E, \forall i \in V \setminus \{1,n\} \quad (3)$$

$$\sum_j x_{1j} = \sum_j x_{jn} = 1, \quad \forall i, j \in V \quad (4)$$

其中约束条件 (2) 和 (3) 共同确保除节点 1 和 n 外的任何一个节点，不存在或只有两条相邻的边，约束条件 (4) 确保节点 1 和 n 是路径的端点。

2 最短路径路由的遗传算法

2.1 量子比特编码与译码

在 QGA 中，采用一种新颖的基于量子比特 (qubit) 的编码方式，即用 qubit 来存储和表达一个基因，该基因可以为“0”态或“1”态，或它们的任意叠加态，即该基因所表达的不再是某一确定的信息，而是包含所有可能的信息，对该基因的任一操作也会同时作用于所有可能的信息。一个由 m 个 qubit 构成的量子染色体可以描述为

$$\begin{bmatrix} \alpha_1 | \alpha_2 | \dots | \alpha_m \\ \beta_1 | \beta_2 | \dots | \beta_m \end{bmatrix}$$

其中 $|\alpha_i|^2 + |\beta_i|^2 = 1$ ($i=1,2,\dots,m$)， $|\alpha_i|^2$ 代表第 i 个基因处于“0”态的概率，而 $|\beta_i|^2$ 代表第 i 个基因处于“1”态的概率。该染色体可以同时表达 2^m 个态的信息。

在求解网络路由的最短路径路由时，量子染色体的量子比特数由网络路由的节点数和所有节点中最大邻接点数共同确定。假设有 n 个节点的网络，所有节点中最大邻接点数为 l ，求得 k 使 $2^{k-1} \leq l \leq 2^k$ ，则量子染色体的量子比特 $m=n \times k$ 。

如图 1 所示网络无向图，则编码时量子染色体的量子比特为 $m=5\times 2=10$ 。

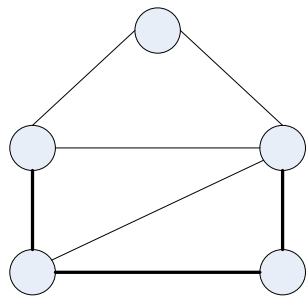


图 1 5 个节点和 7 条边的简单无向图

译码过程需生成矩阵 I 为 $n\times l$ 矩阵， $I(i, j)$ 表示第 i 个节点的第 j 个相邻节点；若与 i 相邻的节点数小于 j ，则 $I(i, j)=0$ 。例如图 1 中，节点 1 相邻的节点为节点 2、3、5，则 $I(1,3)=5$ ， $I(1,4)=0$ 。译码时将量子染色体的量子比特分成 n 组，每组量子比特数为 k ，将第 i 组($i=1,2,\dots,n$)量子比特数看成一个二进制数，并将转换成十进制数 j ，表示第 i 个节点的下一节点为 $I(i, j+1)$ ，然后寻找 $I(i, j+1)$ 的下一节点，直至找到终节点 n 。译码过程中需删除产生无效路径。例如编码(00 01 11 01 00)所表示的路径为(1,2,4,5)。

2

2.2 量子变异策略

量子旋转门调整操作定义如下^[7]：

$$\begin{pmatrix} \alpha'_i \\ \beta'_i \end{pmatrix} = \begin{pmatrix} \cos(\theta_i) & -\sin(\theta_i) \\ \sin(\theta_i) & \cos(\theta_i) \end{pmatrix} \begin{pmatrix} \alpha_i \\ \beta_i \end{pmatrix}$$

(5)

其中， (α_i, β_i) 为量子染色体中第 i 个 qubit 基因， θ_i 为旋转角，由下式决定：

$$\theta_i = s(\alpha_i, \beta_i) \Delta \theta$$

(6)

其中 $s(\alpha_i, \beta_i)$ 和 $\Delta \theta_i$ 分别代表旋转的方向和角度，其值根据表 1 所示的选择策略^[7]确定。

表 1 旋转角选择策略

x_i	b_i	$f(x) > f(b)$	$\Delta \theta_i$	$s(\alpha_i, \beta_i)$			
				$\alpha_i \cdot \beta_i > 0$	$\alpha_i \cdot \beta_i < 0$	$\alpha_i = 0$	$\beta_i = 0$
0	0	False	0	-	-	-	-
0	0	True	0	-	-	-	-
0	1	False	δ	+1	-1	0	± 1
0	1	True	δ	-1	+1	± 1	0
1	0	False	δ	-1	+1	± 1	0
1	0	True	δ	+1	-1	0	± 1
1	1	False	0	-	-	-	-
1	1	True	0	-	-	-	-

表中 x_i 为当前染色体的第 i 位， b_i 为当前最优染色体的第 i 位， $f(x)$ 和 $f(b)$ 分别为当前染色

体和当前最优染色体的适应度函数。 $\Delta\theta_i$ 为旋转角度的大小，当 $x_i=b_i$ 时， $\Delta\theta_i=0$ ；当 $x_i \neq b_i$ 时， $\Delta\theta_i=\delta$ 。 δ 为旋转角度调整的步长， δ 太小将影响算法收敛的速度，太大可能会使结果发散，或早熟收敛到局部最优解。 $s(\alpha_i, \beta_i)$ 为旋转角度的方向，保证算法的收敛。

2.3 算法描述

设第 t 代的量子染色体种群为 $Q(t)=\{q_1^t, q_2^t, \dots, q_n^t\}$ ，其中 n 为种群大小， t 为遗传代数， q_j^t 为定义如下的量子染色体，其长度为 m ：

$$q_j^t = \begin{bmatrix} \alpha_1^t & \alpha_2^t & \dots & \alpha_m^t \\ \beta_1^t & \beta_2^t & \dots & \beta_m^t \end{bmatrix}, j=1,2,\dots,n \quad (7)$$

QGA 其步骤描述如下：

- Step1: 初始化，遗传代数 $t=0$ ，种群 $Q(t)=Q(0)$ ；
- Step2: 由 $Q(t)$ 量子坍缩生成 $P(t)$ ，即对 $Q(t)$ 中各个体实施测量，得到一组状态 $P(t)$ ；
- Step3: 评价 $P(t)$ 中各状态的适应度；
- Step4: 选择并保存最优个体 $b(t)$ 及其适应度值；
- Step5: 停机条件判断：当满足停机条件时，输出当前最优个体，否则继续；
- Step6: 量子变异操作，采用量子旋转门变异策略更新 $Q(t)$ ，得到子代种群 $Q(t+1)$ ；
- Step7: $t=t+1$ ，转到 Step2。

上述算法中，Step1 初始化时， $q_j^t = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & \dots & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & \dots & \frac{1}{\sqrt{2}} \end{bmatrix}, j=1,2,\dots,n$ ；Step2 由 $Q(t)$ 生成

$P(t)$ 的具体测量过程为：随机产生一个属于 $[0,1]$ 的数，若它大于 $|\alpha_i^t|^2$ ，则测量结果取 1，否则取 0。

3 仿真实验与分析

仿真实验是通过该量子遗传算法对已知路由网络求解最短路径路由，并与 Cheng 和 Gen 提出的经典遗传算法^[6]比较。

图 2 比较了采用相同的种群大小时 GA 与 QGA 的收敛特性，横轴表示迭代次数，纵轴表示最短路径时延。由图可知 QGA 能更好地收敛到最短路径路由（迭代次数为 10）；图 3 说明了随着迭代次数的增加，GA 与 QGA 收敛到最优路径的成功率均增加，并且 QGA 的成功率大于 GA 的成功率。

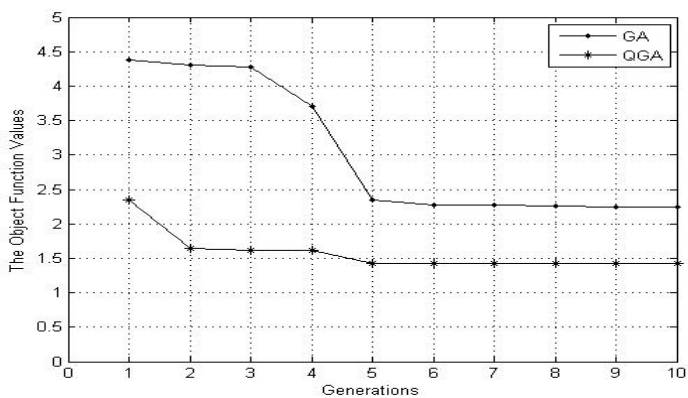


图2 GA 与 QGA 的收敛特性

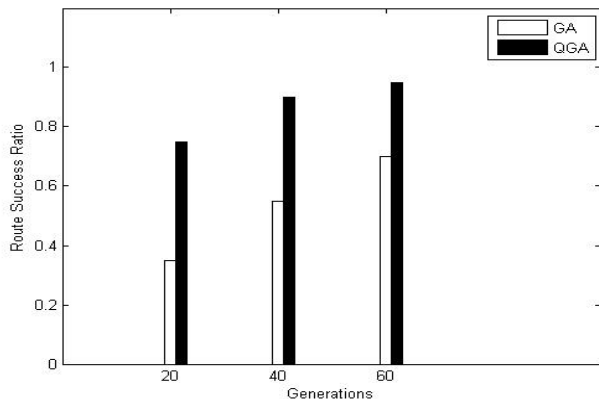


图3 GA 与 QGA 的最优路径选择成功率比较

4 结论

本文主要研究了用量子遗传算法解决最短路径路由的优化问题,提出了一种量子比特编码与译码策略。实验表明,该算法收敛速度快,可靠性高。量子遗传算法由于其并行性和高效性,可以很方便地和现存路由协议结合起来,对通信网中的路由算法进行优化,以获得通信网的最佳路由,例如 OSPF (开放式最短路径优先协议)、BGP (边界网关协议)、RIP (路由信息协议) 等,因此量子遗传算法具有广阔的应用前景。

参 考 文 献

- [1] 孙宝林, 李腊元, 陈 华. 基于遗传算法的最短路径路由优化算法[J]. 计算机工程, 第 31 卷第 6 期, 2005, 3.
- [2] W.Stalling. High-Speed Networks: TCP/IP and ATM Design Principles[M]. Englewood Cliffs, NJ: Prentice-Hall, 1998.
- [3] Holland J.H. Genetic Algorithms and Classifier Systems: Foundations and their Applications [A]. Proceedings of the Second International Conference on Genetic Algorithms, 82-89, 1987

- [4] Han K.-H., Park K.-H., Lee C.-H., Kim J.-H., “Parallel Quantum-inspired Genetic Algorithm for combinatorial Optimization Problem”[J], Proceedings of the 2001 IEEE Congress on Evolutionary Computation, 1422-1429, 2001
- [5] Ahn C W, Ramakrishna R S, A Genetic Algorithm for Shortest Path Routing Problem and the Sizing of Populations[J], IEEE Transactions on Evolutionary Computation, 2002,6(6): 566-579
- [6] 玄光男, 程润伟. 遗传算法与工程优化[M]. 清华大学出版社.2004:258-266
- [7] 李飞. 量子神经网络及其在信号处理中的应用研究[D]. 南京邮电大学博士研究生学位论文. 98-103

作者简介:

刘欣(1986-), 男, 湖北荆州, 南京邮电大学通信与信息工程学院硕士研究生, 主要研究方向: 现代通信中的智能信号处理;

李飞(1966-), 女, 湖南长沙, 南京邮电大学通信与信息工程学院教授, 主要研究方向: 现代通信中的智能信号处理技术、量子信息处理技术;

张杰(1984-), 男, 陕西渭南, 南京邮电大学通信与信息工程学院硕士研究生, 主要研究方向: 现代通信中的智能信号处理。

单载波块传输系统中稀疏信道估计算法研究

黄 帅 朱卫平 孟庆民

(南京邮电大学 信号与信息传输研究院, 江苏 南京 210003)

摘 要: 针对基于循环前缀的单载波块传输系统 (CP-SC), 利用接收信号的自相关函数与信道冲激响应的稀疏特性, 提出了一种新的 CP-SC 稀疏信道估计算法。该算法分为两步完成: 首先, 利用接收信号自相关函数的非零点和稀疏信道中非零抽头系数的位置关系, 确定其非零抽头的位置, 然后采用最小二乘算法对非零抽头系数值进行二次估计, 根据已估计的信道信息, 对接收信号进行 MMSE 均衡处理。仿真结果验证了该算法的有效性。

关键词: 稀疏信道估计; 单载波; 块传输; 循环前缀

Sparse Channel Estimation of Single-Carrier Block Transmission Systems

Huang Shuai Zhu Wei ping Meng Qing min

(Institute of Signal Processing and Transmission, Nanjing University of Posts & Telecommunications, Nanjing 210003, China)

Abstract: In this paper, we propose a new scheme for sparse channel estimation for cyclic prefixed single-carrier block transmission systems (CP-SC), by exploiting the relationship between the locations of non-zero taps and the correlation function of the received symbols. The proposed scheme follows two steps: the first with the non-zero tap detection of the sparse channel; the second with an improved channel estimation utilizing the knowledge of the accurate positions of non-zero taps. Then, CP-SC sparse channel estimation and equalization are presented under least square and MMSE criteria, respectively. The simulation results show that the performance of the new scheme is improved significantly.

Keywords: Sparse Channel Estimation; Single-Carrier; Block Transmission; Cyclic Prefix

1 引言

基于循环前缀的单载波块传输系统 (CP-SC), 又称为单载波频域均衡系统, 具有与 OFDM 相近的复杂度和性能, 但是克服了 OFDM 对于峰均比和频率漂移的敏感性^{[1][2]}。因此, CP-SC 已经成为宽带无线通信中一种具有相当前景的系统技术, 近来受到越来越多的关注。为了充分地利用 CP-SC 系统的潜在性能优势, 准确的信道估计成为 CP-SC 系统设计中的重要环节。

由于信道冲激响应具有稀疏特性，直接使用传统的信道估计算法，如最小二乘估计，估计器的性能比较差^[3]。稀疏信道估计的过程包含两步：首先检测出稀疏信道的非零抽头系数的位置，其次根据非零抽头的准确位置，对非零抽头系数的值进行二次估计。文献[4]研究了接收信号的自相关函数非零点和稀疏信道非零抽头的位置关系。基于这种关系，我们提出了 CP-SC 系统中稀疏信道估计的新方法。

2 系统模型和问题描述

假设单载波块传输模型经过加性噪声干扰的频率选择性信道，如图 1 所示。在发送端，源信号 $s(n)$ 经过串并变换模块处理，得到块 $s(i)$

$$\mathbf{s}(i) = [s(iN), s(iN + 1), \dots, s(iN + N - 1)]^T \in C^N \tag{1}$$

在发送之前，长度为 N 的信号块 $s(i)$ 前面添加长度为 L_{cp} 的循环前缀，用来消除块间干扰 (IBI)。接收端，对信号 $x(n)$ 串并变换后 $\mathbf{x}_{cp}(i)$ ，移除前 L_{cp} 个循环前缀，得到

$$\mathbf{x}(i) = [x(iN), x(iN + 1), \dots, x(iN + N - 1)]^T \in C^N \tag{2}$$

假定在一定数量的 CP-SC 块中，信道系数保持不变，即信道为慢衰落的。该信道可以表示为：

$$h(l) \in C, l = 0, 1, \dots, L - 1 \tag{3}$$

由于循环前缀的长度 L_{cp} 小于信道长度 L ，因此，该系统的输入输出关系可以用矩阵形式进行描述：

$$\mathbf{x}(i) = \mathbf{H}\mathbf{s}(i) + \mathbf{w}(i) \tag{4}$$

其中， $\mathbf{x}(i)$ ， $\mathbf{s}(i)$ ， $\mathbf{w}(i)$ 表示为相应的接收、输入、噪声信号。信道噪声 $\mathbf{w}(i)$ 的定义和输入信号 $\mathbf{s}(i)$ 类似。 \mathbf{H} 是循环矩阵，其第一列是由信道冲激响应 (CIR) 并补上 $N-L$ 个零构成。上述稀疏信道包含很多的零抽头系数，因此信道 \mathbf{h} 中的非零值代表了整个信道的特征。设 p_i 是第 i 个非零抽头的位置， $1 \leq i \leq D$ ， D 是信道冲激响应中非零抽头的数量。因此 \mathbf{h}_{sparse} 作为稀疏信道冲激响应可以表示为

$$\mathbf{h}_{sparse}(n) = \sum_{i=1}^D h(p_i) \delta(n - p_i) \tag{5}$$

本文的目标是利用接收数据的二阶统计量，估计稀疏信道 \mathbf{h}_{sparse} 。

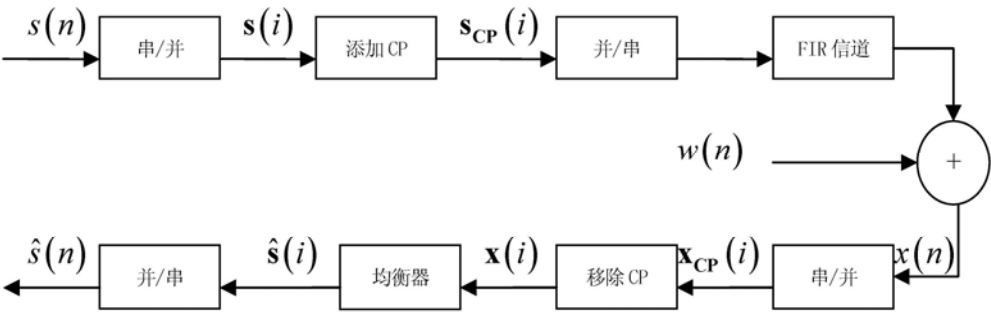


图 1 CP-SC 系统基带模型

3 稀疏信道估计

3.1 基于导频的CP-SC块传输方案

我们将第 i 个发送块的第 n 个信号, 记作 $s(iN+n)$ 。单载波块传输系统发射结构, 如图 2 所示。由于导频位置能极大地影响通信系统的性能, 因此, 我们为接下来的信道估计, 提出了两种不同的导频位置设计方案。

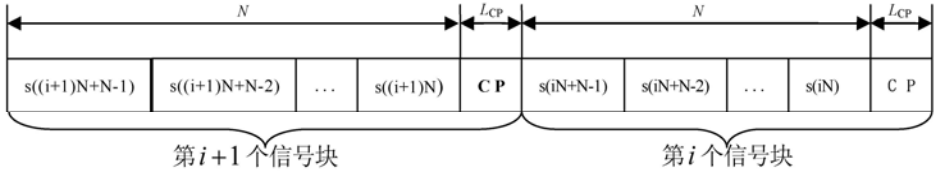


图 2 单载波块传输发射结构

方案一

假定的导频的数量是 K 个, 即从 t 到 $t+K-1$ 。每个 CP-SC 块发送和接收导频可以表示为:

$$\mathbf{s}_t(i) = [s(iN+t), s(iN+t+1), \dots, s(iN+t+K-1)]^T \quad (6)$$

$$\mathbf{x}_t(i) = [x(iN+t), x(iN+t+1), \dots, x(iN+t+K-1)]^T \quad (7)$$

其中, $t \in [0, N-K]$, $i \in \{0, 1, \dots, M-1\}$, M 是 CP-SC 信号块的数量, 在这 M 个块中信道保持不变。需要指出的是, 在本方案中, 导频信号的位置是固定的从 t 到 $t+K-1$, 但是整个导频块可以放置在 CP-SC 块中的不同位置, 随着 t 的变化而变化, 这使得导频插入变得非常方便。

方案二

假定导频放置于每个块的末尾, 即 $t=N-K$

$$\mathbf{s}_t(i) = [s(iN+N-K), \dots, s(iN+N-1)]^T \quad (8)$$

因此导频 $\mathbf{s}_t(i)$ 作为循环前缀 (CP) 的一部分, 进入到 $\mathbf{s}_{cp}(i)$ 的头部。移除 CP 的过程中, 首先分离出混在其中的导频, 这样可利用的导频数量将是方案一的两倍。因此方案二更能提高信道估计的准确性。

3.2 信道估计

文献[4]研究了接收信号的自相关函数非零点和稀疏信道的非零抽头之间的位置关系。首先我们对该位置关系进行简单地分析。接收信号 $x(n)$ 的自相关函数, 在没有噪声的情况下可以表示为:

$$r(l) = E\{x(n)x^*(n-l)\}, l \in \{0, 1, \dots, L-1\} \quad (9)$$

利用卷积的性质, 式 (9) 可以表示为:

$$r(l) = \mathbf{h} \cdot E\{\mathbf{s}_{L-1}(n)\mathbf{s}_{L-1}^H(n-l)\} \cdot \mathbf{h}^H \quad (10)$$

其中,

$$\mathbf{s}_{L-1}(n) = [s(n), s(n-1), \dots, s(n-L+1)]^T$$

根据前面的假设，输入信号的自相关函数 $r(l)$ 可以从式 (10) 中导出：

$$r(l) = \sigma_s^2 \sum_{n=0}^{L-1} h^*(n) \left[\sum_{m=0}^{L-1} h(m) \delta(l+n-m) \right] \quad (11)$$

尽管我们可通过搜寻 $r(l)$ 的非零值来检测非零抽头的位置，但非零抽头的误检测情况依然存在。因此，我们需要更多的信息来降低误检测率，一种可行的方法是利用权值函数 \mathbf{J} 进行估计。

$$\mathbf{J} = \|\mathbf{x}_t(i) - \mathbf{H}\mathbf{s}_t(i)\|^2 \quad (12)$$

对多种可能的非零抽头位置情况，分别计算各自的权值函数 \mathbf{J} ，其中所得结果最小的，则可以认为是较为准确的非零抽头位置。现在，假定非零抽头的位置已经检测出来。对于第 i 个 CP-SC 信号块，我们能够构造出 $(K-L+1) \times L$ 的矩阵称为矩阵 $\mathbf{S}_t(i)$ ，将导频信号按 L 个一组分行逆排列，其中每行之间有 $L-1$ 个导频是重叠的。因此，我们可以得到如下重要的矩阵关系：

$$\mathbf{x}_t = \mathbf{S}_t \mathbf{h}_{\text{sparse}} + \mathbf{w}_t \quad (13)$$

于是，我们可以获得 $\mathbf{h}_{\text{sparse}}$ 的最小二乘估计：

$$\hat{\mathbf{h}}_{\text{sparse}} = (\mathbf{S}_t^H \mathbf{S}_t)^{-1} \mathbf{S}_t^H \mathbf{x}_t \quad (14)$$

3.3 信道均衡

这里，我们采用单载波频域均衡的方法，利用最小均方误差 (MMSE) 准则，对接收信号进行均衡处理，其中，第 i 个接收块可以表示为：

$$\hat{\mathbf{s}}(i) = \mathbf{F}^{-1} \mathbf{W} \mathbf{F} \mathbf{H} \mathbf{s}(i) + \mathbf{F}^{-1} \mathbf{W} \mathbf{F} \mathbf{w}(i) \quad (15)$$

MMSE 均衡器 \mathbf{W}_{MMSE} 可以定义为：

$$\mathbf{W}_{\text{MMSE}} = \arg \min_{\mathbf{W}} E \left[\|\hat{\mathbf{s}}(i) - \mathbf{s}(i)\|^2 \right] \quad (16)$$

对式 (16) 进行代数运算，可以简化为：

$$\mathbf{W}_{\text{MMSE}} = \sigma_s^2 \mathbf{\Lambda}^H \left[\sigma_s^2 \mathbf{\Lambda} \mathbf{\Lambda}^H + \sigma_w^2 \mathbf{I}_N \right]^{-1} \quad (17)$$

因此，对单载波频域均衡系统而言，只需要知道信道的频域响应和高斯噪声过程的方差，就可以设计出一个 MMSE 均衡器。

4 计算机仿真

为了验证该算法的有效性，在单发单收的单载波块传输系统平台上对算法进行了仿真。每一个信号块的长度为 1024 点，循环前缀的长度为 30 点，稀疏瑞利信道模型为包含 3 个非零抽头的 FIR 滤波器，并采用 QPSK 调制。假定非零抽头的位置是 $l_0 = 0$, $l_1 = 6$, $l_2 = 13$ 。

仿真实验一：MSE 与信号噪声比 (SNR)

如图 3 所示，本文提出的两种方案，很显著地提高了稀疏信道的估计性能。特别地，在信噪比为 15dB 时，方案一、二的估计性能增益，分别高出传统方法 4dB 和 6dB。

仿真实验二：误码率（BER）与 SNR

如图 4 所示，在低信噪比（SNR=5dB）的情况下，方案一可以获得 1.2dB 的性能增益，当 SNR 增加到 20dB 时，方案一较之于传统估计方法可以提高 7.3dB。而且，方案二的性能更优于方案一。

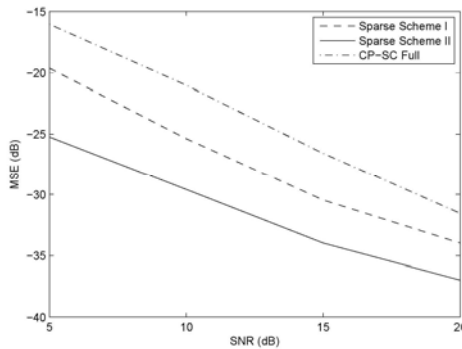


图 3 MSE 与 SNR

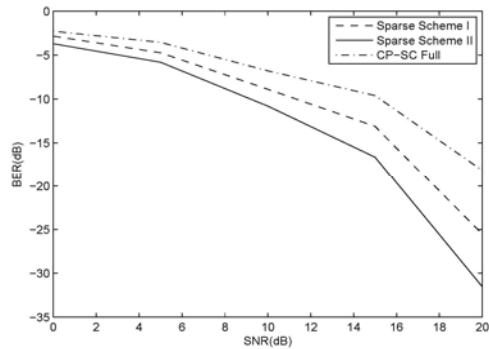


图 4 BER 与 SNR

5 结束语

本文提出了一种 CP-SC 系统中稀疏信道估计的有效算法。该算法利用接收数据的二阶统计量，检测出稀疏信道非零抽头的位置，并在此基础上再进行最小二乘估计，获得非零抽头的数值。一系列计算机仿真结果表明，该算法对于稀疏信道的估计有着良好的性能。

参 考 文 献

[1] Sari H, Karam G, Jeanclaude I. Transmission techniques for digital terrestrial TV broadcasting[J]. IEEE Commun. Mag., vol. 33, no. 2, pp.100-109, Feb. 1995.

[2] Falconer D, Ariyavisitakul SL, Benyamin-Seeyar A, et al. Frequency domain equalization for single-carrier broadband wireless systems[J]. IEEE Commun. Mag., pp. 58-66, Apr. 2002.

[3] Li W, Preisig JC. Estimation of rapidly time-varying sparse channels[J]. IEEE Journal of Oceanic Engineering, vol. 32, no. 4, pp. 927-939, Oct. 2007.

[4] Wan F, Zhu W-P, Swamy MNS. A semi-blind algorithm for most significant tap detection in channel estimation of OFDM systems[C]. Proc. of 2009 IEEE ISCAS, pp. 605-608, May. 2009.

作者简介

黄帅（1984—），男，浙江余姚人，南京邮电大学硕士研究生，研究方向为无线通信与信号处理技术。

朱卫平（1962—），男，江苏如皋人，南京邮电大学教授，博士生导师，研究方向：超三代(B3G)无线移动通信中的关键信号处理技术等。

孟庆民（1965—），男，江苏滨海人，南京邮电大学讲师，硕士生导师，现主要研究方向为无线 mesh 网协同策略设计、MIMO 网络和宽带多载波与多址系统性能评估等。

基于无扰动的最大速率传输MIMO系统半盲信道估计^{*}

陈 钟 朱卫平

(南京邮电大学, 信号处理与传输研究院, 江苏 南京, 210003)

摘 要: 提出一种 MIMO 平坦衰落信道下的无扰动半盲信道估计方法。其思想是对传输信号自相关矩阵进行特征分解。通过发送少量的自相关矩阵的信息到接收端, 去除扰动误差从而提高了信道估计的性能并减少导频数目实现最大速率传输。计算机仿真显示该方法在估计性能方面, 优于传统最小均方估计方法以及闭式半盲估计方法。

关键词: 多输入多输出; 最大速率传输; 扰动误差; 信道估计;

Base Perturbation Free Semi-Blind Channel Estimation Scheme for MRT MIMO Systems

Chen Zhong Zhu Wei Ping

(Institute of Signal Processing and Transmission, Nanjing University of Posts and Telecommunications, Nanjing, China 210003)

Abstract: In this paper, propose a MIMO flat fading channels without perturbation semi-blind channel estimation method. The idea is to transmit the signal autocorrelation matrix eigen-decomposition. By sending a small amount of self-correlation matrix of information to the receiver, remove the perturbation error so as to enhance the performance of channel estimation and reduce the number of maximum rate of pilot transmission. Computer simulation shows that the method in estimating the performance, superior to the traditional least mean square estimation methods, and closed semi-blind estimation method.

Keywords: MIMO; MRT; signal-perturbation-free; channel estimation;

1 引言

近年来, MIMO 通信系统由于不仅可以通过获得复用增益来提高系统容量, 而且能够利用分集增益来抵抗信道衰落, 从而成为研究的热点。为了达到这些性能效果, 准确的信道估计对于 MIMO 系统是至关重要的。最大速率传输策略就是预编码和成形矩阵就是对应于信道矩阵 \mathbf{H} 的左, 右主要特征向量 $\mathbf{v}_1, \mathbf{u}_1$ 。发送端 \mathbf{v}_1 对信号进行预编码, 经过 MIMO 系统发送给接收端, 接收端则利用 \mathbf{u}_1 对接收信号进行处理, 从而得到发送端信号。所以该系统性能效果

^{*} 国家自然科学基金项目 (60872104), 江苏高校自然科学基金项目 (08KJD510001), 教育部博士基金项目 (20080293004)

就取决于发送端对信道的了解程度^[1]。

在频率平坦 MIMO 下的信道估计的算法中,常见的有基于导频的方法和半盲估计的方法。在论文[2]-[4]论述了,以下两种方法。一种是基于导频传统最小均方估计算法 (CLSE),该方法是利用导频,从而得到信道矩阵,对信道矩阵进行奇异值分解来得到 $\mathbf{v}_I, \mathbf{u}_I$ 。另一种是闭式半盲算法 (CFSB),该方法通过使用盲算法从信号中估计 \mathbf{u}_I ,再利用导频,从中估计出 \mathbf{v}_I 。

由于 CFSB 方法中存在信号扰动误差^{[5][6]}。利用论文[7]-[8]中的无扰动分析,提出了一种新的无信号扰动的发射方案 (SPF-CFSB),从而提高了 CFSB 方法的估计性能。该方法利用对发送信号自相关矩阵进行特征分解,设计了一种有效的发送方案来减少接收端得信号扰动误差。结果显示新方法比原来的 CFSB 具有更好的性能效果。

2 简要回顾基于导频和半盲信道估计算法

2.1 系统模型数学表示

对于一个 $(N_r \times N_t)$ MIMO 系统,信道经过一个平坦,非频率选择性瑞利衰落。信道特性表示为一个 $N_r \times N_t$ 的矩阵 $\mathbf{H}=[h_{ij}]$ 是一个具有零均值的方差为 1 的独立同分布复高斯随机变量。 h_{ij} 表示从第 j 个发送天线到第 i 个接收天线的信道增益。信道矩阵的奇异值分解可以表示为 $\mathbf{H}=\mathbf{U}\mathbf{\Sigma}\mathbf{V}^H$ 。令 $\mathbf{v}_I, \mathbf{u}_I$ 分别表示 \mathbf{V} 和 \mathbf{U} 的第一列。发送信号: $\mathbf{s}(n)=[s_1(n), \dots, s_{N_t}(n)]^T$,其各个元素是服从独立同分布,均值为零且 $\delta_s^2=1$ 的高斯随机变量。接收信号: $\mathbf{x}(n)=[x_1(n), \dots, x_{N_r}(n)]^T$ 。可表示为:

$$\mathbf{x}(n) = \mathbf{H}\mathbf{s}(n) + \mathbf{w}(n) \tag{1}$$

其中,噪声矢量: $\mathbf{w}(n)=[w_1(n), \dots, w_{N_r}(n)]^T$,元素之间不相关且方差 $\delta_w^2=1$ 。在每一个数据块,前 L_p 个符号用于导频信号。令 L_d 表示为传输的白信号(未经过预编码的信号)的符号个数。也就是,总计 $L_p+L_d=N$ 符号先于发送预编码信号。由于 L_d 个白信号携带信号,所以不造成带宽的浪费。

2.2 传统最小均方估计算法 (CLSE)

当使用正交导频序列 \mathbf{S}_p ,最小均方信道估计 \mathbf{H}_e ,首先得到:

$$\mathbf{H}_e = \mathbf{X}_p \mathbf{S}_p^\dagger = \frac{1}{\gamma_p} \mathbf{X}_p \mathbf{S}_p^H \tag{2}$$

\mathbf{X}_p 是接收导频序列。根据最大似然估计的特性,对信道矩阵 \mathbf{H} 的最大似然估计 \mathbf{H}_e ,进行奇异值分解,得到 \mathbf{v}_I 和 \mathbf{u}_I 的最大似然估计。也就得到了发送与接收端得预编码和成形矢量符号。

2.3 闭式半盲估计算法 (CFSB)

假定发送信号为白信号, \mathbf{u}_I 的最大似然估计为接收信号自相关矩阵 \mathbf{R}_x 的主要特征值。通过对 \mathbf{R}_x 进行特征分解,可以计算得到 \mathbf{u}_I 的估计:

$$\mathbf{U}\tilde{\mathbf{\Sigma}}^2\mathbf{U}^H = \mathbf{R}_x \tag{3}$$

u_e 表示为 u_I 的估计，也就是从盲估计得到的 U 的第一列。

利用以下的限制最小均方问题的解^{[3][4]}:

$$\hat{v}_e = \arg \min_{v \in C^{N_t}, \|v\|=1} \left\| \tilde{X}_p - v^H \tilde{S}_p \right\|^2 \quad (4)$$

在假设已知 u_I 条件下，可以估计 v_I 。

$$\hat{v}_e = \frac{S_p X_p^H u_I}{\|S_p X_p^H u_I\|} \quad (5)$$

利用导频序列 S_p ，将 u_e 代替 u_I 代入(5)，从而估计出 v_I 。

根据以上的分析，首先从(3)中得到估计 u_e 。然后利用 L_p 个导频符号 S_p ，将 u_e 代替 u_I 代入(5)从而得到 v_I 的估计。

3 无扰动发射方案(SPF-CFSB)

由于采用有限长度的数据来计算自相关矩阵，即使在没有噪声存在的条件下，也存在着信号扰动。接收信号的自相关矩阵 R_x ，包含信号扰动，可写为：

$$R_x = E[x(n)x^H(n)] - \delta_w^2 I = H[I + \Delta R_s]H^H + \Delta R_w \quad (6)$$

在无噪声情况，所有由噪声引入的扰动噪声将消去。于是，(6) 将写成：

$$\hat{R}_x = H[I + \Delta R_s]H^H \quad (7)$$

首先，先将 ΔR_s 进行特征分解。由于 ΔR_s 是一个自共轭矩阵，所以 ΔR_s 的特征值分解可写成：

$$\Delta R_s = \sum_{i=1}^{N_t} \sigma_i u_i u_i^H \quad (8)$$

其中， $\sigma_i (i=1,2,\dots,N_t)$ 是取值范围为 $(-\infty, \infty)$ 的特征值， $u_i (i=1,2,\dots,N_t)$ 是其相应的特征向量。

接下来，基于 (8)，将 ΔR_s 进一步分解，使得 ΔR_s 的信息更容易传给接收端，从而去除信号扰动误差。

$$\Delta R_s = \eta S_{per} S_{per}^H \quad (9)$$

其中， η 是个比例因子， S_{per} 是包含 ΔR_s 信息的矩阵。

将 S_{per} 分成两部分 S_{pos} 和 S_{neg}

$$S_{per} = S_{pos} - S_{neg} \quad (10)$$

整个 S_{pos} 由如下表示：

$$S_{pos} = [S_{pos,1} S_{pos,2} \cdots S_{pos,L_{pos}}] \quad (11)$$

其中： $S_{pos,i}$ 满足以下等式：

$$\eta S_{pos,i} S_{pos,i}^H = \sigma_{pos,i} u_{pos,i} u_{pos,i}^H \quad (12)$$

同理， S_{neg} 根据如下表示：

$$S_{neg} = [S_{neg,1} S_{neg,2} \cdots S_{neg,L_{neg}}] \quad (13)$$

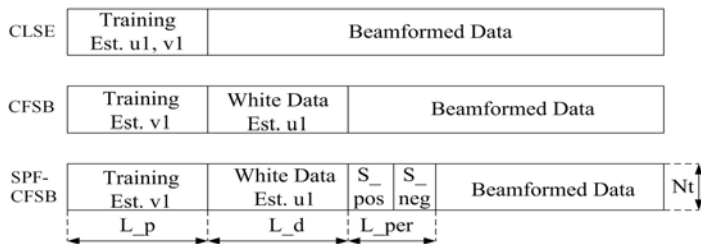


图 1 各种发射方案比较

从以上讨论，所提出的传输结构，如图 1，包括导频信号，白信号，以及附加信号 S_{pos} 和 S_{neg} 。通常， η 的选择根据发送天线的个数和数据的长度。根据仿真显示， $\eta=1$ 时，在 4×4 MIMO 系统中，可以得到很好的估计效果。在该系统中，传输扰动信息只需要 4 个时隙。而白信号的时隙个数为 100。

4 实验结果

此节采用计算机仿真来验证各种算法的性能。仿真建立在一个瑞利衰落信道下，系统具有 4 个发送天线和 4 个接收天线。数据采用 16-QAM 调制方式。

本实验采用的准则是，奇异值之间的测量误差准则：

$$\|v_1 - \hat{v}_1\|_{DN}^2 \stackrel{\Delta}{=} 2(1 - |v_1^H \hat{v}_1|) \quad (14)$$

满足：

$$\|v_1 - \hat{v}_1\|_{DN}^2 = \min_{\varphi \in (-\pi, \pi]} \|v_1 - \hat{v}_1 e^{j\varphi}\|^2 \quad (15)$$

实验 1: MSE - SNR

在第一个实验中，根据不同的信噪比下所得到的估计 v_l 误差准则性能。仿真是基于 10000 个蒙特卡罗次数，其中 10 个时隙用于导频，50 个用作白信号。图 2 显示出根据 CLSE, CFSB, SPF-CFSB 所得到得曲线。可以看出在信噪比小于 18dB 时，SPF-CFSB 是其中性能最好的。然而，在信噪比大于 18dB，传统的 CLSE 方法就显示出其的优势。而且，SPF-CFSB 在任何信噪比下都比 CFSB 多 4 个 dB，从而说明 SPF-CFSB 相对于 CFSB 的优势。

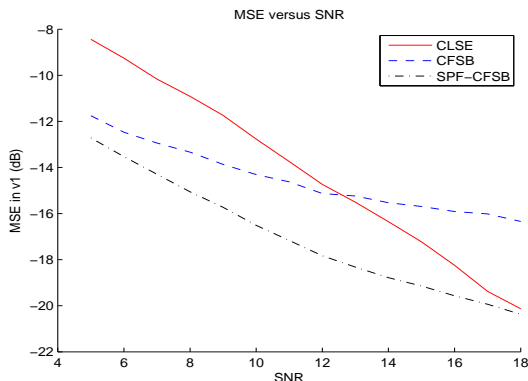


图 2 MSE - SNR

实验 2: MSE - Pilot Length

该实验根据不同的导频长度来获得各算法信道估计性能。图 3 显示的是在 10000 个蒙特卡罗次数下, 传输信噪比为 10dB 情况下的性能曲线。可以发现在导频长度低于 23 时, 采用 CFSB 的性能优于 CLSE 方法。然而在导频增加的情况下, CLSE 就优于 CFSB。在不同导频长度中, SPF-CFSB 的性能是最好的。例如, 在导频长度为 58 时, SPF-CFSB 方法的性能仍高于 CFSB 方法 3dB, 从而隐含 SPF-CFSB 不只在信噪比为 10dB 时, 效果比 CFSB 好。

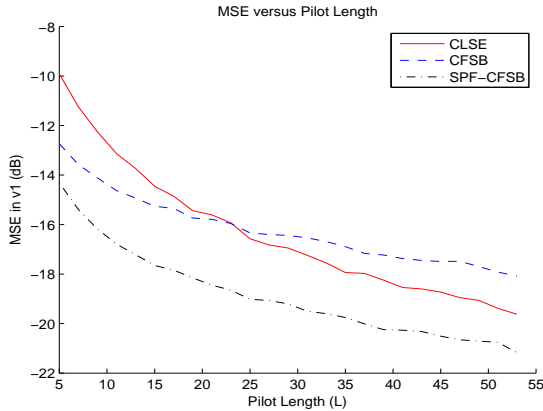


图 3 MSE - Pilot Length

5 结论

本文先简述传统最小均方估计算法 (CLSE) 和闭式半盲估计算法 (CFSB) 的信道估计方法。接着提出了新的无扰动发射方案 (SPF-CFSB)。计算机仿真结果表明在 MIMO 平坦衰落信道下, SPF-CFSB 在不同导频和不同信噪比下都优于 CLSE 和 CFSB 算法。

参 考 文 献

- [1] T. K. Y. Lo, "Maximum ratio transmission," *IEEE Trans. Commun.*, vol.47, no.10, pp.1458-1461, Oct. 1999.
- [2] D. Pal, "Fractionally spaced equalization of multipath channels: a semi-blind approach," in *Proc. Int. Conf. Acoustics, Speech, Signal Processing (ICASSP)*, vol. 3, pp. 9-12, 1993.
- [3] A. K. Jaganatham, C. R. Murthy, and B. D. Rao, "A semi-blind MIMO channel estimation scheme for MRT," *Proc. Int. Conf. Acoustics, Speech, Signal Processing (ICASSP)*, vol.3, Philadelphia, PA, Mar. 2005, pp. 585-588.
- [4] C. R. Murthy, A. K. Jaganatham, and B. D. Rao, "Training-based and semiblind channel estimation for MIMO systems with maximum ratio transmission," *IEEE Trans. on Signal Processing*, vol. 54, no.7, pp. 2546-2558, Jul. 2006.
- [5] Z. Xu, "Perturbation analysis for subspace decomposition with applications in subspace-based algorithms," *IEEE Trans. On Signal Processing*, vol. 50, no. 11, pp. 2820-2830, Nov. 2002.
- [6] F. Wan, W.-P. Zhu, and M. N. S. Swamy, "Perturbation analysis of subspace-based semi-blind MIMO channel estimation approaches," *IEEE International Symposium on Circuits and Systems (ISCAS)*, 2008, pp. 129-132.
- [7] F. Wan, W.-P. Zhu, and M. N. S. Swamy, "A signal perturbation free transmit scheme for MIMO channel estimation," in *Proc. IEEE Vehicular Technology Conference (VTC-Fall)*, Sept, 2008, pp. 1-5.

- [8] F. Wan, W.-P. Zhu, and M. N. S. Swamy, "A signal perturbation free whitening-rotation-based semi-blind approach for MIMO channel estimation," *IEEE Trans. on Signal Processing*, vol. 57, no. 8, pp. 3154-3166, Aug. 2009.

作者介绍:

陈钟（1985—），男，福建连江人，南京邮电大学硕士研究生，研究方向为无线通信与信号处理技术。

朱卫平（1963—），男，江苏如皋人，南京邮电大学教授，博导，研究方向为无线通信与网络信号处理，现代通信中的智能信号与信息处理等。

数字集群系统中移动链路管理的研究与实现

张 建 何志敏

(南京邮电大学信号处理与传输研究院 江苏 南京 210003)

摘 要: 对数字集群通信系统中终端侧的移动链路管理、控制进行了研究。对网络层较低子层移动链路实体层 (MLE) 及逻辑链路层 (LLC) 进行了功能描述, 协议分析, 并依据功能提出实现基本链路建立、移动性管理控制及信令传输的流程设计。最后, 在嵌入式微处理器硬件平台上, 利用有限状态机及消息机制实现终端侧协议功能的软件设计。

关键词: 移动链路管理; 逻辑控制; 数字集群; 协议实现

Research and Implementation of Mobile Link Management in Digital Trunked System

Zhang Jian He Zhi min

(Institute of Signal Processing and Transmission, Nanjing University of Posts & Telecommunications, Nanjing 210003, P. R. China)

Abstract: This paper deals with the research of mobile link management and control on the terminal side in digital trunked communication system. It also describes function of the sub-layer of the network layer Mobile Link Entity (MLE) and the Logical Link Layer (LLC), analyzes the protocol process and proposes the flow of the establishment of basic link, the control of mobile management and the signaling transmission based on the function. Finally, we achieve the protocol design of the terminal by means of finite state machine and message mechanism on the embedded microprocessor hardware platform.

Keywords: mobile link management; logical control; digital trunked; protocol implementation

0 引言

集群系统属于专用移动通信系统, 与公众的蜂窝移动系统相比具有呼叫接续快、可调度, 群组内用户共享前向信道、半双工通信方式和支持私密呼叫和群组呼叫等特点, 在专业通信领域发挥着巨大的作用。TETRA 数字集群通信系统是 ETSI (欧洲通信标准协会) 针对专用移动通信的需要而设计的基于数字时分多址技术的具有统一标准开放性专业移动通信系统, 集快速呼叫建立、组呼、直通模式运作、分组数据、电路模式数据传送、高安全性等多项功能于一身, 应用相当广泛。

TETRA 数字集群空中接口协议是按照 OSI 七层参考模型定义的, 协议栈对应七层模型的下三层, 分别为物理层, 数据链路层以及网络层。本文重点研究集群通信系统中空中接口协

议层中的终端侧对无线链路管理与控制，衔接了协议栈的第三层网络层较低子层(移动链路实体 MLE)与第二层数据链路层的较高子层（逻辑链路控制 LLC），并在嵌入式处理器平台上编程实现终端侧的协议功能。

1 移动链路实体MLE与逻辑链路控制LLC协议分析

1.1 功能描述

在 TETRA 数字集群空中接口协议栈中,移动链路控制实体(MLE)子层位于第三层网络层,负责移动台终端到基站之间链路的建立与维护，登记区之内的移动性管理、移动台身份管理和服务质量控制等功能。MLE 子层同时具有高层会聚的功能,为第三层的实体提供一个访问第二层实体服务的统一平台。

而逻辑链路控制 LLC 是位于移动链路控制实体 MLE 的下层协议，主要功能包括传输数据的确认业务与非确认业务（在基本链路业务中也可使用 FCS 的可选扩展差错控制），当需要传送大量数据(分组数据传输)或要求更好的业务时,就要使用高级链路,此时可提供以下业务：下行链路中点对多点的非确认消息传输、窗口机构、长消息的顺序分段、点对（多）点传送的选择性重传等。

1.2 移动链路管理信令流程

通过研究及对比当前成熟 GSM 信令流程方案，我们可以将 MLE 复杂的移动链路管理流程分为基站覆盖小区链路扫描选择，正服务小区链路测量，相邻小区链路监视并分级，以及小区漫游切换。处理从管理实体到较低层的本地管理信息 MLE 和 LLC 协议栈基本链路之间使用原语进行信令通信。原语分为四种类型:请求原语、指示原语、响应原语以及证实原语。本地的 MLE 使用请求原语向 LLC 请求一项业务，而 LLC 使用证实原语向 MLE 证实与请求业务相关的操作已完成。对方 LLC 使用指示原语通知 MLE 操作结果。MLE 使用响应原语来确认接收来自 LLC 的指示原语。

a) 小区链路初选阶段，在收到上层移动管理 MM 激活请求后，MLE 处理空口信道选择，扫描可用链路（由 MAC 层进行逻辑信道到物理信道的转换），满足标准选择即可选用可用链路后，并在此链路接受基站广播。此时向 MM 报告，MM 发起注册，采用确认链路（非确认则 LLC 应发送两次消息），LLC 需等待对端响应 BL-ACK，此时交换侧 MM 发出注册成功（或失败响应），终端移动台侧 MM 向 MLE 发出更新消息，MLE 通知 CMCE 呼叫控制实体可以打开此时就建立了一条信令链路。

b) 正服务小区质量测量及相邻小区监视分级，在一个小区进行注册后，即开始对当前链路质量不断进行检测，并根据收到的网络广播包（LLC 采用非确认链路传送方式含有相邻小区参数信息）MLE 发给下层（LLC 透传给 MAC）对相邻小区链路质量参数进行检测，并按照链路质量参数进行相邻小区分级，此时可得到一个优先小区，为迁移漫游做好准备。

c) 小区漫游。管理小区释放、改进和可用无线准则，包括了漫游通告(Announcements)和宣告(Declarations)管理。小区再选过程应该确保 MS 选择一个小区，在其中能有效解码下行链路数据，并有高可靠性的上行链路通信，最低条件应满足路径损耗参数 $C1>0$ ，并且在电路模式呼叫中 MS 不超过最大路径延时。

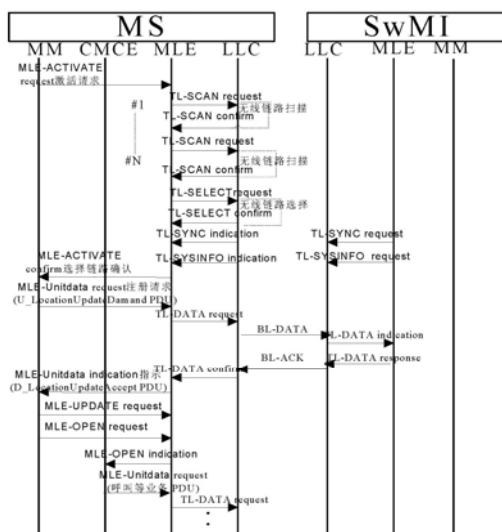


图1 小区初选流程（MS：移动台终端，SwMI：交换和管理基础设施）

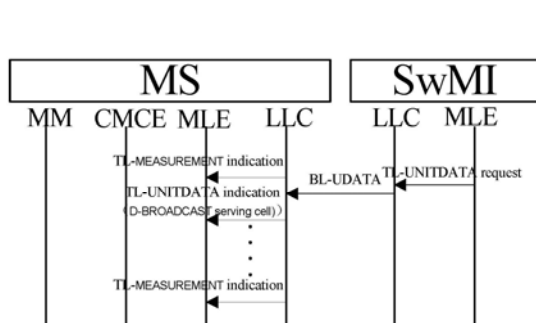


图2 正服务小区监视

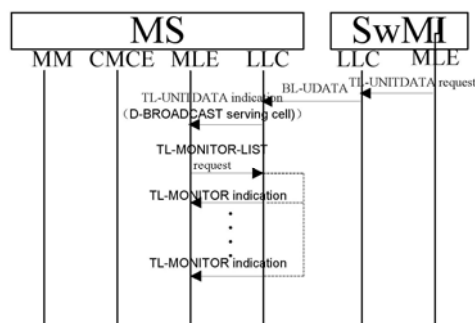


图3 相邻小区监测

i. 无宣告小区再选：当 MLE 无连接时（即链路完全未建立），MS-MLE 无论对正服务小区或新小区都不能告知小区改变。向较高的第 3 层实体 MLE-BREAK 指示原语，指示 C 面无效。如果已有分级过后的优先相邻小区，则选择该小区，否则开始扫描，选择，注册等小区初选过程，完成后向上发出 MLE-REOPEN 指示原语，呼叫链路控制重新打开。

ii. 无通告小区再选：与当前小区链路中断，采用无通告小区再选。过程类似于无宣告小区再选，但重新选择完小区后，使用 MLE-RESUME 指示原语告诉 CMCE 可以恢复已存在的任何电路模式呼叫，CMCE 会往对端发送呼叫恢复 PDU，MLE 发射一个 U-RESTORE PDU，利用 LLC 确认业务该 PDU 包含 CMCE 呼叫恢复 SDU。如果恢复呼叫成功，MLE 应该从 SwMI 中接受一个 D-RESTORE ACK PDU，MLE 应该给 CMCE 发射 MLE-RESTORE 证实原语，包含 CMCE 恢复信令 PDU。

iii. 通告类型小区再选：分为 3 种类型。在不知道新小区的情况下（未收到正服务小区含相邻小区信息的广播 PDU）通告类型 3 小区再选向旧 SwMI 发射一个 U-PREPARE PDU，准备小区再选，交换设备回送 D-NEW CELL PDU，之后处理过程类似于无通告小区再选。通告类型 2 是在已知需转换的新小区信息并将向旧 SwMI 发射一个 U-PREPARE PDU（含新小区 ID），等待 D-NEW CELL PDU，此时旧小区无法直接转换信道或者无法发出有关新小区的网

络信息。D-NEW CELL PDU 中信道指配命令应为立即由终端切换信道，此时终端即开始选择新信道链路（不需扫描）。通告类型 1 小区再选是在已知需转换的新小区业务信道信息（TCH）并将向旧 SwMI 发射一个 U-PREPARE PDU（含注册信息，并且含有新小区 ID）收到新小区的 D-NEW CELL PDU，含注册成功 PDU 信道指配命令应为 MAC 信道改变，这样“无缝”地实现小区再选，无需发送恢复呼叫 PDU。在小区再选过程中遇到失败情况（注册，恢复，转移等），则从分级小区列表中选择另一优先小区。以通告类型 2 小区链路再选为例：

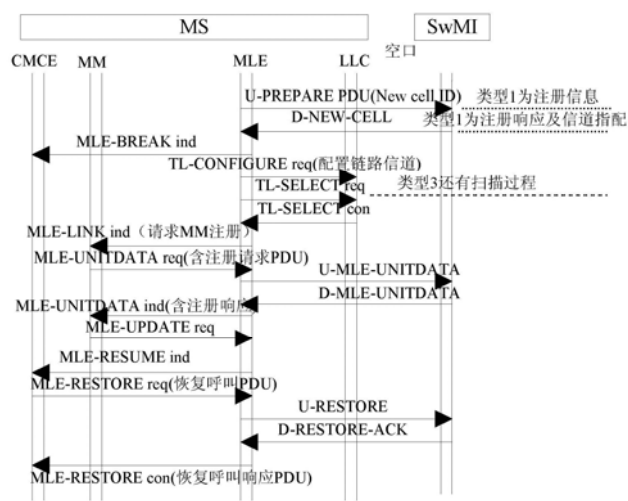


图 4 通告类型 2 小区再选

2 MLE与LLC结构设计

我们可以定义 MLE 完成的四个主要功能：附着管理、数据传输、网络广播、管理功能。与此相对应定义了四个实体：附着管理实体（Attachment）包括对小区初选和小区再选程序。数据传输实体(Data Transfer)有如下几种工作方式：基础 LLC 业务的选择,通知上层实体 CMCE 和 SNDP 禁止和允许接入通信资源，第 2 层业务端点路由和复用（包括增加和清除 MLE 协议控制信息），MLE SAP 和其他 MLE 功能实体的路由和复用，映射的业务质量（如优先权、吞吐量、传输业务）。网络广播功能实体(Network Broadcast)：对于终端侧网络信息的接收和分析，同时用同步和系统信息广播配置第 2 层 MAC。管理功能实体(Management Entity)：处理网络管理过程。

而对于 LLC 层子系统的结构主要指的是空口协议栈软件结构，该软件结构可分为两大部分：基本链路业务部分和高级链路业务部分。基本链路部分包括 4 个模块，分别是：逻辑链路模块（LOGIC）、基本链路确认业务模块（BASA）、基本链路非确认业务模块（BASU）、基本链路格式化模块（Formatter）。高级链路部分包括 5 个模块，分别是：逻辑链路模块（LOGIC）、高级链路发送模块（TXA）、高级链路确认接收模块（RXA）、高级链路非确认接收模块（RXU）、高级链路格式化模块（Formatter）。

实现基本的语音数据及短消息通信，基本链路确认（非确认）业务模块又以下几个模块组成：初始化函数、基本链路确认（非确认）PDU 生成函数、基本链路确认（非确认）PDU 编码函数、LLC 部分消息的拆封函数、内部消息的生成函数、优先先权队列链表、全局参

数的管理函数。基本链路格式化模块包括：初始化函数、PDU 解码子函数、模块路由函数（发往确认或非确认模块）、消息的拆封函数。逻辑链路模块的软件结构包括：初始化函数、消息的路由转发功能。

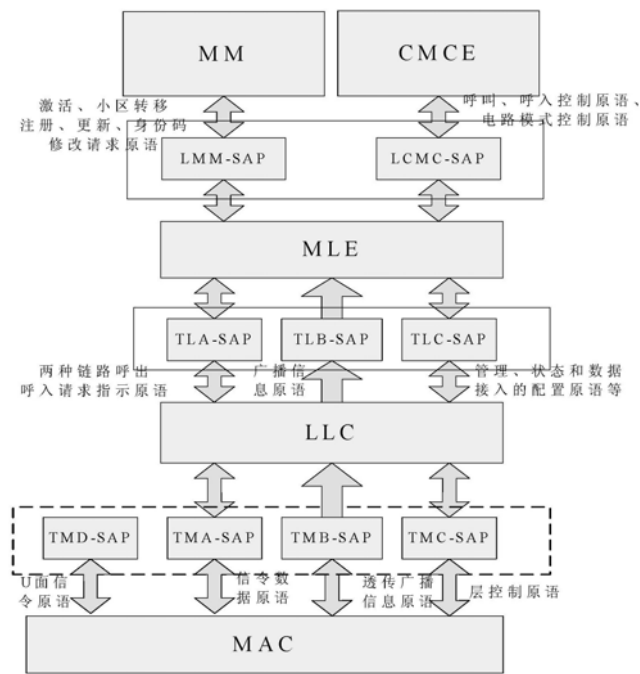


图 5 协议栈层间结构

我们定义 MLE 和 LLC 层与层之间的接口为 SAP 点结构，存放消息 ID。

3 基于流程方案的软件实现

我们采用有限状态机及消息队列机制的设计方法，对 MLE 及 LLC 的信令流程进行处理。消息 ID 定义：对应消息名称唯一宏定义消息 ID。

缓冲区定义：采用短整型的一维数组。每个缓冲区定义一组读写指针，采用 FIFO 方式处理消息。

定义模块处理状态：针对不同模块定义不同状态，可具体为模块状态类型：{空闲，状态 1，状态 2...}。

定义消息格式：每个不同的消息原语对应不同的消息类型，可具体为消息类型：{信息元素（字段）1，信息元素（字段）2...}，然后根据消息类型定义唯一消息全局变量。

程序实现：我们将 MLE 和 LLC 处理函数放入一个 ARM 侧子进程中，1）调用 MLE 及 LLC 初始化函数 2）不断循环调用 MLE 和 LLC 处理函数，采用轮询的方法，当有原语消息产生，即可调用了不同的子模块处理函数 3）双核通信过程 LLC 从 DSP 接收 MAC 层发上来的消息或向 MAC 发送消息（一组乒乓过程不会同时产生）。具体地对于每个子模块可以形成一个处理函数：

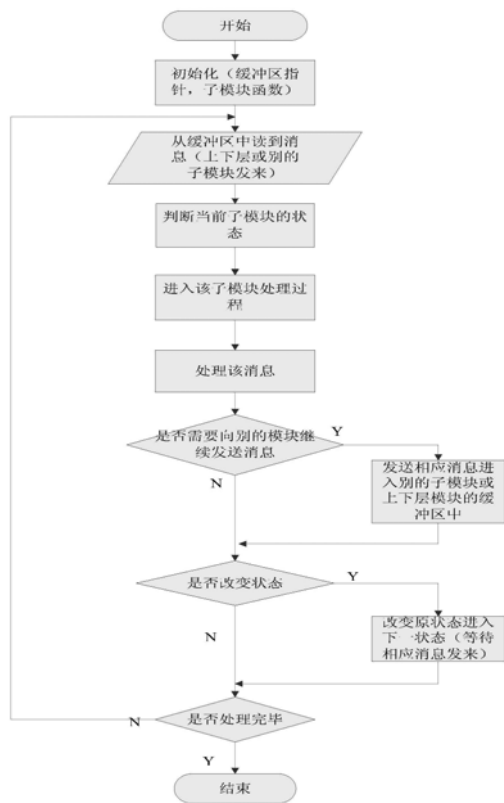


图 6 程序处理流程

1) 从当前模块缓冲区中读取消息 ID

2) switch(消息){

case 消息ID1: if(当前状态==状态1)

处理1; (消息内容为全局变量)

else if(当前状态==状态2) 处理2;

...

break;

case 消息ID2: if(当前状态==状态1)

处理3;

else if(当前状态==状态2) 处理4...

...

break;

...

default:break;}

4 硬件实现

我们采用的硬件平台是由 TI 公司生产的双核(ARM926 和 DSP5510)处理器 OMAP5912, MLE 及 LLC 及上层程序协议程序加载在 ARM926 (嵌入式 Linux 平台) 单元中。为方便程序

能够在 Windows 平台下测试, PC 机经过仿真器 (TDS560) 对加载在目标板中的代码, 并在 CCS 软件平台上进行实时测试。目标板先后通过串口和射频口与基站及交换机通信。而在协议层内部 LLC 侧与下层 MAC 层通过双核通信机制进行通信, 短消息利用 mailbox 机制, 而对于长消息则利用 mailbox 传递消息 ID, MMU 传递消息内容, ARM 处理器将控制 DSP 的 MMU 部分地址映射到一块相同的空间。

我们利用硬件软件平台测试协议实现, 并按照协议定义流程与对端进行乒乓过程。经测试完成小区初选, 基本链路数据传输, 小区监测, 小区漫游(实际测试中多为选择相邻质量较好的小区)等的通信过程。

5 结束语

目前, 我国数字集群还属于新兴产业, 专用数字集群移动通信设备的商业化几乎还是空白。因此, 大力发展我国自主标准的数字集群系统, 提供更高性能价格比的数字集群系统, 是十分必要的。本文针对 TETRA 数字集群通信系统, 介绍了系统协议中 MLE 及 LLC 的功能, 根据现有通信过程研究其协议流程, 并设计实现 MLE, LLC 协议程序, 完成无线链路控制及管理, 为进一步研究数字集群通信系统及其信令链路处理协议打下了良好的基础。

参 考 文 献

- [1] 郑祖辉, 陆锦华, 郑兰.数字集群移动通信系统[M] (第 2 版) [M].北京: 电子工业出版社, 2005.
- [2] 胡宏林, 徐景, 3GPP LTE 无线链路关键技术[M]., 北京: 电子工业出版社, 2008.
- [3] Richard Dewey, Remi Sfez 。 DESIGN OF THE TETRA MOBILE RADIO AIR INTERFACE PROTOCOL[J].IEE Conference Publication, n 387, 1993, p 38-43.
- [4] ETSI EN 300 392-2 Terrestrial Trunked Radio Voice Plus Data, Part2[S].

作者简介:

张建 (1985.06), 男, 籍贯:重庆, 南京邮电大学信号与信息处理专业 2007 年硕士, 研究方向: 无线通信与信号处理。联系方式: TEL: 13770692184; Email:zhangjian_ssm@hotmail.com

何志敏 (1984.12), 女, 籍贯: 江西, 南京邮电大学信号与信息处理专业 2007 年硕士, 研究方向: 无线通信与信号处理。联系方式: TEL: 15895883035;Y070605@njupt.edu.cn

数字无中心移动通信系统MAC层的协议设计

强文炜

(南京邮电大学 信号处理与传输研究院, 江苏 南京 210003)

摘要: 在通信系统全面数字化的趋势下, 专网通信的全面数字化已经迫在眉睫。对欧洲电信标准化协会下的 TETRA (Terrestrial Trunked Radio)协议标准下的 DMO (直通模式) 进行了全面研究, 简要介绍了数字无中心多信道选址移动通信的原理。在 DSP 集成开发环境 CCS2.0 平台上, 设计并实现了一套无中心系统的 MAC 层协议模型, 详细阐述了 MAC 层的帧结构与突发结构以及逻辑信道与 MAC 层帧的映射关系。设计模块已经成功嵌入实验室数字无中心移动通信系统中, 运行效果良好, 充分证明了模块的可行性。

关键词: 数字无中心; 移动通信; MAC 层; 帧结构

Design of Protocol of MAC Layer in Digital Land Mobile Communication System without Central Controller

Qiang Wen wei

(Institute of Signal processing and Transmissing, Nanjing University of Posts and Telecommunications, Jiangsu Nanjing, 210003, China)

Abstract: Under the tendency to communication system being digitized overall, it is apparently that general digital in the special-communication is overwhelming. Considering the DMO (Direct Mode Operation) of TETRA (Terrestrial Trunked Radio) protocol standard of ETSI, the theory of land mobile communication system requirements using multi-channel access techniques without central controller is introduced. Based on the platform of integrated development environment of CCS2.0, a layer protocol model which is suitable for the land mobile communication system without centre controller is designed. Then the Frame Structure and burst Structure of MAC layer and mapping between MAC frame and logical channel are described in detail. It is embedded in the system of the lab's system, which has been well implemented. Its feasibility is proved in experiment.

Keywords: digital centre less; mobile communication; MAC layer; frame structure

0 引言

鉴于移动通信技术的发展以及灾后救援等复杂环境下对于通信系统的特殊要求, 无中心多信道选址移动通信系统由于其没有交换控制中心、组网灵活, 具有自组织、自愈功能, 成为移动综合通信系统的重要组成部分。因此对该系统的研发应该加以重视。MAC(Medium Access Control)层的主要功能在于根据呼叫控制层信息生成相应的 MAC 层帧结构, 提供信道

编解码、差错控制、流量控制等功能。以语音数据传输为例，详细介绍了无中心多信道选址移动通信系统中 MAC 层的帧、物理层突发结构以及 MAC 帧与逻辑信道的映射关系。

1 数字无中心系统

1.1 数字无中心系统简介

无中心多信道选址移动通信系统是一种不采用交换控制中心的集中控制，而由各移动分别设定无线通信链路的分散控制方式的自集群系统。本文研究的无中心移动通信系统是集语音、数据、短信息、文本、图片及态势传输、终端显示等功能，具有通信指挥一体的综合通信系统。该系统采用先进合理的数字化技术，实现在同一对线路上进行数字化、远距离、综合业务的分布式自组织、自愈网络；具有频率自适应和功率自适应特点，其子网在不小于同时传输 4 路、6 路、8 路话音情况下，在被复线上的通信距离 分别不小于 10 公里、8 公里和 6 公里；可通过中继、网关实现语音、数据等多业务双工通信，融入移动综合通信网。

1.2 无中心多信道选址通信的原理

无中心多信道选址移动通信系统，首先选取频点最低的一个信道作为信令信道，用于发送呼叫信令和广播信息。在用户开机后，所有用户驻留在信令信道上等待接收广播消息。通过接收广播消息获得当前系统的定时信息以及业务信道的忙闲信息和其他用户的忙闲信息，由于本系统是采用 TDMA(时分多址)方式，且没有交换控制中心，需要有用户充当主控用户，在固定的帧号和时隙号上发送复帧号、帧号、时隙号，实现网同步，与当前系统达到同步。在与系统同步上以后，空闲用户驻留在信令信道上时刻监听是否有信令呼叫自己。

若有一用户欲与另一用户通话，首先需要查询当前业务信道忙闲表与用户忙闲表，看欲呼叫的用户是否空闲。若被叫用户空闲，则主叫向被叫用户发送呼叫建立信令。被叫根据主叫发送的呼叫建立信令利用测距算法测算出与主叫用户之间的距离，根据该距离和当前业务信道的忙闲信息确定是否需要中继，然后选择合适的业务信道号与发送、接收时隙号和业务帧发送能量，将上述信息包含在确认帧中回送给主叫用户。主叫用户收到该确认后，发送振铃信息，在接到被叫摘机的信令后，主叫与被叫用户切换至事先约定好的业务信道号和发送接收时隙号上，按照之前设定的语音帧发送能量发送语音帧。待通话结束，释放信道，呼叫双方重新回到信令信道守候。

与有交换控制中心通信系统不同的是，无中心多信道选址系统没有固定的交换控制中心。由于是 TDMA(时分多址)系统，同步的方案在无中心多信道选址通信系统中很关键。在本系统中采用采用后发起者跟随先发起者的跟随同步方式。所有在线用户跟随信令信道同步。即系统中始终存在一个主控用户，该用户在每个复帧的固定帧号、时隙号上发送当前系统的复帧号、帧号与时隙号，所有在线用户需要以主控用户发送的定时信息为参照，更新自身的同步计时器。即使已经进入通话状态，处于业务信道上的用户，也需要在该固定的同步帧到来的时间段之前切换回信令信道上接收定时信息。否则会认为该用户已经进入失步状态，进入到失步处理流程。

2 MAC层的通信协议

本文研究的无中心多信道选址通信系统由于需要适用于灾后救援等复杂通信环境下，需要同时满足通信与指挥的功能。根据该特殊环境下的功能和需求，本系统很多功能类似于数字集群调度通信系统。因此本系统参照了欧洲较成熟的 ETSI（欧洲电信标准化协会）制定的 TETRA（Terrestrial Trunked Radio)协议标准下的 DMO（Direct Mode Operation）协议标准。该标准中规范了无线接口的物理层和数据链路层及业务和设备的数据设置。^[1]

2.1 MAC层系统协议结构

TETRA 标准中 DMO（直通模式）的三层协议构架是由 OSI(Open System Internetwork) 的 7 层模型演化而来，继承了通用的分层结构，被通信分层框架所接受。其中媒体接入控制（MAC）为第 2 层，主要处理各个用户对介质的共享。其中层 3 为 DMCC(Direct Mode Call Control)，为呼叫控制层，主要用于处理信令交互过程；层 2 为 DLL（Data Link Layer），为数据链路层，主要实现 MAC 层的功能；层 1 为 PL（Physical Layer），为物理层。其整体协议结构框架如图 1 所示^[2]，在 MAC 层，协议栈被垂直的分为 2 个部分，一是 User Plane（用户面），用于传送信息(如语音、U 面信令、文件等)，但是无寻址能力；另一个用于信令的 Control Plane（控制面），带有寻址能力。呼叫控制层 DMCC（Direct Mode Call Control）位于第

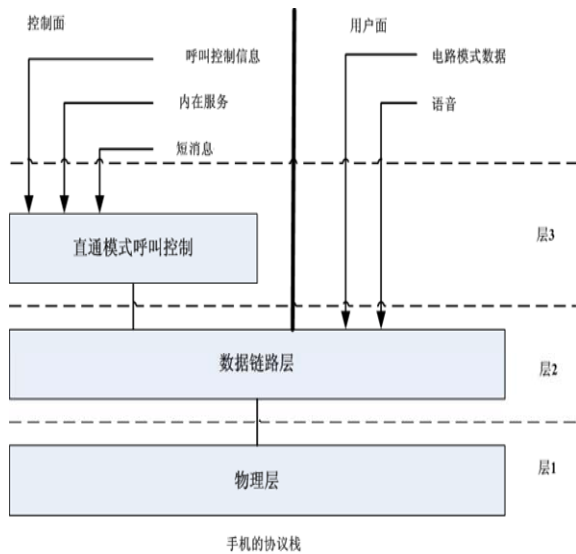


图 1 协议栈结构

3 层，和控制面相连，负责呼叫控制（如寻址等）以及信令的解析，提供无线电终端支持的一般业务，如单呼、群呼、会议、组呼、短信等，以及增强业务。如具备调度系统特点的强拆、强插等补充业务。数字无中心系统不仅可以承载语音传输，完成普通的通话对讲要求；同时还能承载数据业务，完成诸如短消息发送、文件和图片传输等传统模拟对讲机系统无法完成的任务。数据传输是模拟系统和数字系统本质的分界线，数据传输业务的实现使得专网

移动通信系统向数字化通信又迈进了一大步。

2.2 MAC层的主要工作任务

本系统中 MAC 层主要分为上、下 MAC 层，分别处理不同的功能。图 2 是本模块的整体框架。

其中上 MAC 层主要负责以下任务：

- 1) 与 DMCC 层进行信令和用户数据的交互，完成帧中各项协议数据的配置；
- 2) 广播帧信息的生成与接收，进入失步流程后的处理；
- 3) 业务信道忙闲表与用户忙闲表的管理；
- 4) 业务信道、时隙的管理；

下 MAC 层主要负责以下任务：

- 1) 信道编解码：CRC(Cyclic Redundancy Check)、卷积码、交织、扰码等；
 - 2) 数据编解码模块：根据任务队列传来的数据包完成信令的组合解析及信道编解码任务；
- 主要任务包括信令和数据的组合分解、CRC、卷积码、交织、扰码等算法的编解码实现；

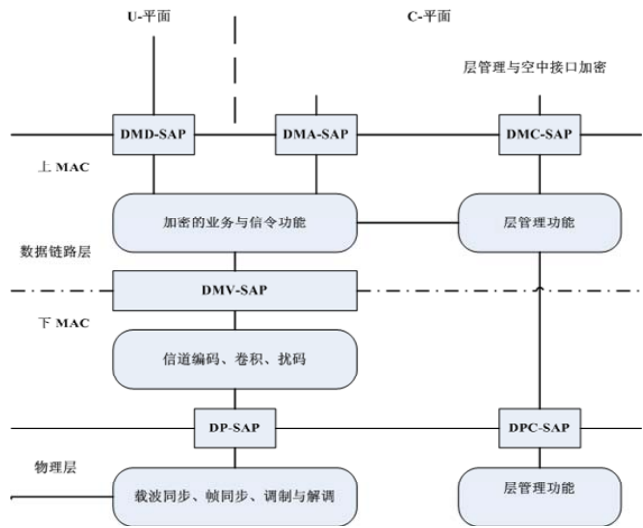


图 2 MAC 层模块整体框架

3) 传输与接收模块：主要负责数据的发送和接收控制管理，是链路层和物理层的接口。发送时，按照任务队列中任务的要求，把编码过的信令和数据封装成帧，并置位发送准备标志位，通知物理层发送数据；接收时，按匹配到的同步字把接收到的数据分成帧，每接收完一个帧，置位接收标志位，通知编解码模块提取信令和数据信息；

3 MAC层帧结构与突发结构

3.1 MAC层帧传输结构与逻辑信道的映射^[3]

MAC 层中数据传输时被封装成帧，主要有 2 个目的：①由帧头和帧尾定界一个数据传输

单位，使接收方正确识别一次数据传输的开始和结束；②可以通过添加复杂的帧头来增强协议的功能，为系统扩展留下了空间。在本系统中 MAC 层主要的传输单位也是帧。DMCC 层以及用户面数据通过 SAP(服务接入点)进入到上 MAC 层，上 MAC 层根据该信令不同的功能分别封装成 MAC 层的帧，在上下 MAC 层中间的 SAP 点-DMV SAP 点映射到不同的逻辑信道上。

MAC 层在逻辑信道上提供数据传送业务，逻辑信道类型集合是为 MAC 层提供的不同类型的数据传输业务而定义的。逻辑信道通常可以分为两类：控制信道和业务信道。控制信道用于传输控制平面信息，而业务信道用于传输用户平面信息。

表 1 是 MAC 层帧与逻辑信道的映射以及所承载的不同功能的信息（PDU 是指协议数据单元）其中 SCH-S、BNCH、SCH-F 用于承载控制平面信息，而 TCH-S、SCH-H 用于承载用户平面信息：

表 1 MAC 层帧与逻辑信道的映射

MAC 层帧类型	MAC 层 PDU 类型	逻辑信道类型
信令帧	DMAC-SIGNAL	SCH-S
广播帧	DMAC-SYNC	BNCH
业务帧	语音	TCH-S
短信帧（帧头部分）	DMAC-DATA	SCH-S
短信帧（分片）	DMAC-FRAG	scH-F
U 面信令帧	DMAC-U-SIGNAL	SCH-H

以 DMAC-SIGNAL 为例，该 MAC 层帧内容最丰富，主要用来承载呼叫建立、信道协商等信令。下面简单介绍一下 MAC 层帧的域以及其中各参数所代表的含义：

```
typedef struct
{
MACPDUTypeType    MACPDUType;
unsigned int       DestinationAddress;
Boolean           SourceAddressInd;
unsigned int       SourceAddress;
unsigned int       LengthIndicator;
MessageTypeType    MessageType;
unsigned int       Chan_Id;
unsigned int       TxSlotId;
unsigned int       RxSlotId;
unsigned int       TxEnergy;
unsigned char      DM_SDU[3];
}DMAC_SIGNALType;
```

下面对结构体中的元素做相应解释：

MACPDUType: MAC 层 PDU 的类型，分别表示该 PDU 承载的内容

DestinationAddress: 目标地址

SourceAddressInd: 源地址指示，该参数为 1 表示需要在 PDU 中加入源地址

SourceAddress: 源地址

LengthIndicator: 长度指示, 表示当前 PDU 所占的长度, 以 Bit 为单位

MessageType: 消息类型, 用于指示从呼叫控制层 (DMCC) 传送过来的层 3 PDU 类型

Chan_Id: 选择的用于传送语音的业务信道号

TxSlotId: 发送时隙号, 一帧中用于发送业务信息的时隙

RxSlotId: 接收时隙号, 一帧中用于接收业务信息的时隙

TxEnergy: 发送能量等级

DM_SDU: 呼叫控制层实际包含的信息, 该信息 MAC 层不可见, 接收端接到后将 SDU 直接通过通信原语传送至 DMCC 层

从发送端来看, 在完成 MAC 层帧与逻辑信道的映射后, MAC 层帧进入到下 MAC 层的信道编码过程, 不同的逻辑信道对应于不同的编码方式。

3.2 物理层突发结构

MAC 的突发结构, 突发是物理信道的最小单元。MAC 层帧需要生成相应的物理层突发, 被实际的物理信道所承载, 每个突发由载波同步序列、训练序列以及块所组成。其中载波同步序列是物理层用于判断一帧数据的到来, 一般为全 1; 不同训练序列用于区分不同的信道, 如区分信令信道和业务信道; 块承载的是实际的 MAC 层信息。在完成 MAC 层帧的编码后, 分别对应信令信道与业务信道, 下 MAC 层要生成 2 种突发结构: DSB(信令突发)与 DNB (正常突发)。

其中 DSB 用于承载信令信道和广播信道的信息, 对应于 MAC 层帧即为 DMAC-SIGNAL、DMAC-SYNC、DMAC-DATA、DMAC-FRAG。DNB 用于承载业务信道上的信息, 对应于 MAC 层帧中的语音帧及 DMAC-U-SIGNAL。图 3 与图 4 分别是 DSB 与 DNB 两种突发的结构: 信令信道突发结构 DSB 如下:

12bit 载波同步	22bit 训练序列 (有 2 种训练序列, 分别表示广播和信令)	块 228bit
------------	--------------------------------------	----------

图 3 DSB 突发结构

业务信道突发结构 DNB 如下:

32bit 载波同步	38bit 训练序列	块 608bit
------------	------------	----------

图 4 DNB 突发结构

其中块中存放的就是经过 CRC、卷积码、扰码交织过后的 MAC 层帧。

在接收端, 下 MAC 层在接收到物理层传送上来的突发后, 根据训练序列判断是信令信道还是业务信道的信息。在经过去交织、去扰码、维特比译码、去 CRC 等操作后, 根据 MAC 帧帧头中的信息, 解码出 MAC 层 PDU 的类型, 并将 SDU (服务数据单元) 传送给上 MAC 层。上 MAC 层在解析得到自身有用的信息后, 将 MAC 层 PDU(协议数据单元)中的信息传送给 DMCC (Direct Mode Call Control) 层, 由其进行信令流程的操作。

4 结束语

系统中介绍的 MAC 层帧结构、突发结构满足了无中心多信道选址移动通信系统在复杂通信环境下实现集语音、数据、短信息、文本、图片及态势传输、终端显示等功能，具有通指一体等功能的要求。本系统模块已嵌入实验室无中心多信道选址移动通信系统，其中语音、短信息等功能已经过验证，效果良好。

参 考 文 献

- [1] 郑祖辉, 陆锦华, 郑兰. 数字集群移动通信系统[M] (第 2 版) [M]. 北京: 电子工业出版社, 2005.
- [2] 薛鸿媛, 谢晓明. 数字无中心通信系统数据链路层的设计与实现[J]. 计算机工程与设计, 2008, (9): 4654-4655
- [3] ETSI EN 300 396-3 Terrestrial Trunked Radio(TETRA): Technical Requirements for Direct Mode Operation(DMO); Part 3: Mobile Station to Mobile Station(MS-MS) Air Interface(AI) Protocol [S]: 56-158

作者简介:

强文炜 (1982.10), 男, 籍贯: 江苏, 南京邮电大学信号与信息处理专业 2007 年硕士, 研究方向: 无线通信与信号处理。联系方式: TEL: 13914719869; Email: qiangwenwei@126.com

基于 802.16e 交织编码部分的 FPGA 设计与实现

毛智强 雷维嘉 谢显中

(重庆邮电大学个人通信研究所 重庆 400065)

摘要: 给出了一种适用于 802.16e 系统的自适应交织方案, 在 RS-CC 级联码的基础上, 交织器可以根据信道的变化采用不同的交织算法对数据进行处理, 从而适应不同的调制编码方案。方案采用 VHDL 语言实现, 并通过 Quartus7.2 软件仿真和链路联调验证, 结果表明可以将突发错误离散成随机错误, 解决 RS-CC 级联码对于突发错误的纠错能力有限问题。

关键词: 802.16e; 交织; VHDL; 突发错误

Design and Implementation of 802.16e Interleaved Coding/Decoding by Using FPGA

Mao Zhi qiang Lei Wei jia Xie Xian zhong

(Institute of Personal Communication of Chongqing University of Posts and Telecommunications, Chongqing 400065, P.R. China)

Abstract: This paper introduces an interleaving scheme in 802.16e system. According to the changes of the channel, it can use relevant interleaving algorithm to deal with its data in order to adapt to different encodes/decodes and modulation/demodulation schemes. The simulation in Quartus7.2 software and the debugging in the link of Wimax show that this scheme languaged by VHDL can transmit burst errors into random errors to improve the correcting capacity of RS-CC coding.

Keywords: 802.16e; interleaving; VHDL; burst errors

1 引言

802.16e 作为一个新兴的无线城域网技术标准, 比 802.11 可以在更广阔的地域范围内提供“最后一英里”的宽带接入以及更好的移动性。因此, 作为 4G 可能的技术演进方向之一, 802.16e 技术的设计与实现将非常具有现实意义。

本文给出了一种适用于 802.16e 系统的自适应交织方案, 在 RS-CC 级联码的基础上, 交织器可以根据信道的变化采用不同的交织算法对数据进行处理, 从而适应不同的调制编码方案。本方案采用 VHDL 语言实现, 并通过 Quartus7.2 软件仿真和链路联调验证, 结果表明可以将突发错误离散成随机错误, 解决 RS-CC 级联码对于突发错误的纠错能力有限问题。

2 802.16e交织技术的基本原理

在 802.16e 标准中使用的交织结构属于分组结构。所有被编码的数据都应该被一个块交织器交织，交织器的大小为 N_{cbps} 。交织器由两部分组成(如图 1 所示)：第一部分是载波之间的交织，确保相邻的编码比特被映射到非相邻的子载波，防止某个子载波遭受深度衰落而导致的突发性的错误；第二部分交织是确保邻近的比特被交替地映射到星座图中较低或较高的比特。

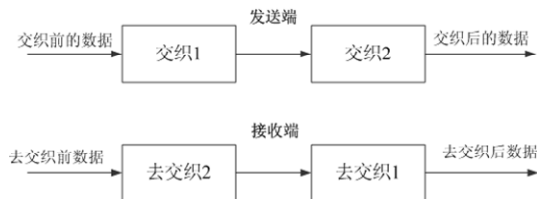


图 1 802.16e 交织器组成

在这里，我们令 N_{cpc} 为每个子载波中被编码的比特数,比如 BPSK,QPSK, 16-QAM 和 64-QAM 分别为 1,2,4 和 6 令 $s = \text{ceil}(N_{\text{cpc}} / 2)$ ，（ceil 表示向上取整，例如 $\text{ceil}(3.2) = 4$ ）， k 为第一次排列前被编码比特的顺序号； m_k 为第一次排列后第二次排列前的顺序号； j_k 为第二次排列后调制映射前的顺序号。

第一次排列由式 1 定义：

$$m = (N_{\text{cbps}} / 12) \cdot k_{\text{mod}(12)} + \text{floor}(k / 12) \quad k = 0, 1, \dots, N_{\text{cbps}} - 1 \tag{1}$$

$k_{\text{mod}(12)}$ 表示对 K 取 12 的模运算，floor 表示向下取整。

第二次排列由式 2 定义：

$$j = s \cdot \text{floor}(m_k / s) + (m_k + N_{\text{cbps}} - \text{floor}(12 \cdot m_k / N_{\text{cbps}}))_{\text{mod}(s)} \quad m = 0, 1, \dots, N_{\text{cbps}} - 1 \tag{2}$$

同样，去交织器也由两次排列组成，令 j 为在第一次排列前所接收比特的顺序号， m_j 为第一次排列后第二次排列前的顺序号， k_j 为第二次排列后卷积解码前的顺序号。

第一次排列由式 3 决定：

$$m_j = s \cdot \text{floor}(j / s) + (j + \text{floor}(12 \cdot j / N_{\text{cbps}}))_{\text{mod}(s)} \quad j = 0, 1, \dots, N_{\text{cbps}} - 1 \tag{3}$$

第二次排列由式 4 决定：

$$k_j = 12 \cdot m_j - (N_{\text{cbps}} - 1) \cdot \text{floor}(12 \cdot j / N_{\text{cbps}}) \quad j = 0, 1, \dots, N_{\text{cbps}} - 1 \tag{4}$$

表 1 显示了对于各种调制和子信道数，交织块的大小：

表 1 各种调制编码方式对应交织块的大小

	16 个子信道	8 个子信道	4 个子信道	2 个子信道	1 个子信道
	N_{cbps}				
BPSK	192	96	48	24	12
QPSK	384	192	96	48	24
16-QAM	768	384	192	96	48
64-QAM	1152	576	288	144	72

本文采用的是基于 OFDM 的 IEEE 802.16e 物理层技术,在 OFDM 调制模式下,基于 RS-CC 的级联编码的交织编码。表 1 所列出的调制方式和子信道数指的是数据在经过 RS-CC 的级联编码、交织之后 Wimax 系统所采用的调制方式和子信道个数 N,图 2 为 OFDM 系统的基本模型框图, S/P 表示串并转换, P/S 为并串转换。

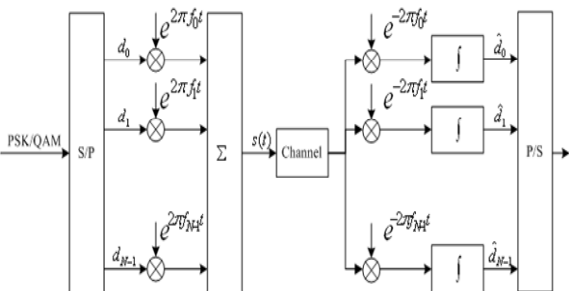


图 2 OFDM 系统基本模型图

一个 OFDM 符号是多个经过 PSK 或者 QAM 调制的子载波信号之和。这里, N 表示子信道的个数, T 表示 OFDM 符号的宽度, $d_i (i=0,1,\cdots,N-1)$ 是分配给每个子信道的数据符号, 第 i 个子载波的频率为 $f_i=i/T+f_c (i=0,1,\cdots,N-1)$, f_c 是第 0 个子载波的载波频率, $\text{rect}(t)=1, |t|\leq T/2$ 。

以用 16-QAM 调制、16 个子信道的情况为例, 交织 1 矩阵数据的写入和读出顺序如图 3 所示, 输入数据在时钟的作用下依次被放在第一行的 0、1、2...62、63, 第二行的 0、1、2...62、63, 直到最后一行被放满; 被放满的数据在接下来的一个时钟沿开始依次从第一列的 0、1...11、12, 第二列的 0、1...11、12, 直到最后一列的数据位读出并交给交织 2 矩阵。

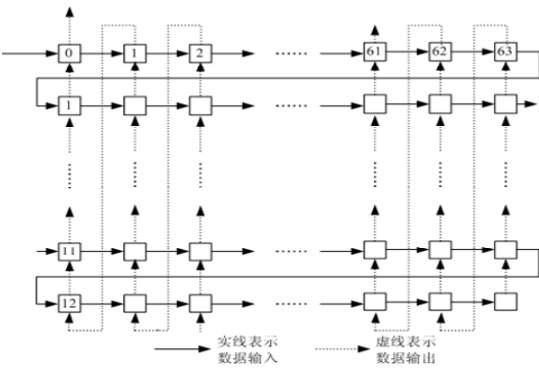


图 3 交织 1 矩阵

交织 2 矩阵数据的写入和读出顺序如图 4 所示, 输入数据在时钟的作用下依次被放在第一行的 0、1、2...10、11, 第二行的 0、1、2...10、11; 放满后, 数据在接下来的一个时钟沿开始依次从位置 0、13、2、15、4、17、6、19、8、21、10、23、12、1、14、3、16、5、18、7、20、9、22、11 被读出送到串并转换器。

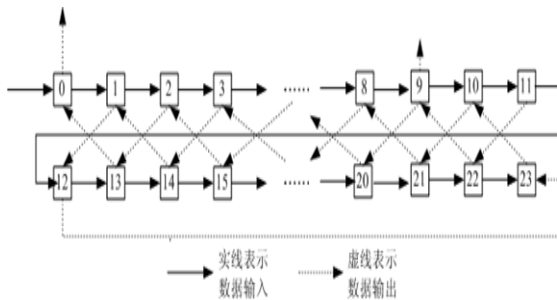


图 4 交织 2 矩阵

3 802.16e 交织技术的FPGA实现

对于各种调制编码方式，我们需要选择不同的交织器来满足自适应的要求。而对于不同的交织器而言，区别只是交织块的不同，在电路中表现为读写地址的顺序不同。因此，可以得到如下设计思路：一片 RAM 做随机存储器，一个交织地址生成器用来产生交织地址，由一个计数器产生顺序地址。当需要顺序地址时，选择器选择计数器的输出送至 RAM 地址线；当需要交织地址时，选择器选择从交织地址生成器产生的交织地址送至 RAM 地址线，由此可得框图如图 5 所示。

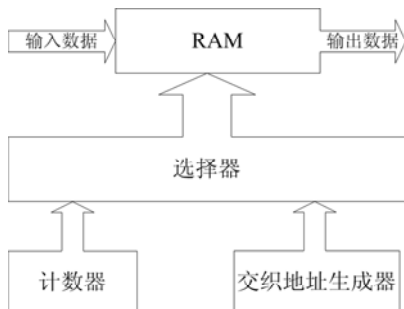


图 5 自适应交织器实现框图

交织时，计数器产生的顺序地址分别作为 RAM 的写入地址，交织地址生成器产生的交织地址作为 RAM 的读出地址；去交织时，计数器产生的顺序地址作为 RAM 的读出地址，从交织地址生成器的输出的交织地址作为 RAM 的写入地址。

为了能够让交织器能够连续的处理接收到的数据，本文采用了 2 片 RAM 来实现乒乓操作，当一片 RAM 接收顺序输入数据的时候，另一块 RAM 输出交织后的数据；当接收完一组数据的同时也刚好发送完前一组数据，此时交换 2 片 RAM 的工作方式，然后重复上述操作，实现对数据的连续处理，实现框图如图 6 所示。

而对于交织地址编码器的设计，通过式 1、式 2 计算和查表 1 交织块大小，我们可以得到表 2 中各种对应算法。交织地址编码器根据选择模式的不同在表 2 中选择相应的算法，最终实现自适应的交织地址生成器。

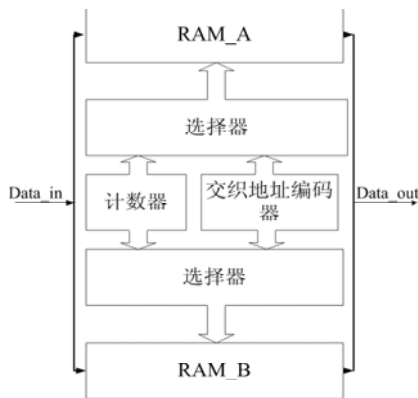


图6 能够连续处理接收到的数据的自适应交织器实现框图

表2 各种调制编码方式对应的交织地址算法

		16 个子信道	8 个子信道	4 个子信道	2 个子信道	1 个子信道
BPSK	交织 1	前一交织地址加 16 取 191 的模	前一交织地址加 8 取 91 的模	前一交织地址加 4 取 47 的模	前一交织地址加 2 取 23 的模	顺序地址
	交织 2	无				
QPSK	交织 1	前一交织地址加 32 取 383 的模	前一交织地址加 16 取 191 的模	前一交织地址加 8 取 95 的模	前一交织地址加 4 取 47 的模	前一交织地址加 2 取 23 的模
	交织 2	无				
16-QAM	交织 1	前一交织地址加 64 取 167 的模	前一交织地址加 32 取 383 的模	前一交织地址加 16 取 191 的模	前一交织地址加 8 取 95 的模	前一交织地址加 4 取 47 的模
	交织 2	顺序地址中的奇数地址加 12 取 23 的模，0 和偶数地址不变				
64-QAM	交织 1	前一交织地址加 96 取 1151 的模	前一交织地址加 48 取 575 的模	前一交织地址加 24 取 287 的模	前一交织地址加 12 取 143 的模	前一交织地址加 6 取 71 的模
	交织 2	顺序地址中的第 1 个地址（即 0 和 3 的倍数）不变，第二个地址加 24 取 35 的模，第三个地址加 12 取 35 的模，如此循环				

在这里，对于表 2 的算法我们还可以归纳如下：

（1）对于交织 1 的交织地址我们只需将前一交织地址加上 $N_{\text{cpc}} N_c$ （ N_{cpc} 为子载波中被编码的比特数，比如 BPSK、QPSK、16-QAM 和 64-QAM 分别为 1, 2, 4 和 6， N_c 为子信道个数），之后再取对应交织块大小减 1 的模（交织块大小见表 1），就得到了交织 1 地址；

（2）对于交织 2，BPSK 和 QPSK 没有交织 2，而不同子信道数的 16-QAM 的交织 2 地址算法一致（算法见表 2），不同子信道数的 64-QAM 的交织 2 地址算法也一致（算法见表 2）。

本文用 VHDL 语言编程实现了基于表 1 的各种参数的自适应交织，并通过 Quartus7.2 软件仿真，链路联调验证。图 7 是交织器的状态转移图。

如图 6 所示，交织器共有 3 个状态：R_IDLE（空闲），R_STAT A（工作状态 A）和 R_STAT B（工作状态 B），状态转移过程如下。

（1）R_IDLE 状态：当交织器被复位信号复位后，状态即将立刻进入这一状态。在这个状态下，交织器一直在等待一个数据帧起始信号 FRAM_S。当 FRAM_S=‘1’，交织器读取交织参数，设定交织地址生成器的工作模式，初始化计数器（置零），状态转为 R_STAT A（工作状态 A）；

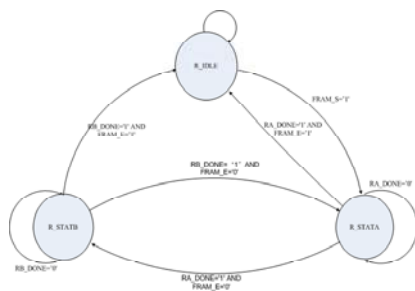


图 7 自适应交织器状态转移图

(2) **R_STATA 状态**: 在这个状态下, 交织器的计数器每加 1, 就把待交织数据写入 **RAM_A**、计数器输出顺序地址置入 **RAM_A** 的写入地址, 交织地址生成器输出交织地址置入 **RAM_B** 的读出地址。当标志位 **RA_DONE** 变成 ‘1’ 并且 **FRAM_E** 不等于 ‘1’ 时, 即 **RAM_A** 写入, **RAM_B** 输出完成而一帧没有结束时, 初始化计数器 (置零), 交织器转入下一状态 **R_STATB**。若标志位 **RA_DONE** 变成 ‘1’ 并且 **FRAM_E** 等于 ‘1’ 时, 即 **RAM_A** 写入, **RAM_B** 输出完成而一帧结束时, 交织器转入下一状态 **R_IDLE**;

(3) **R_STATB 状态**: 工作流程与 **R_STATA** 状态类似。

去交织器的状态转移图同交织器的状态转移图类似, 区别主要是去交织器用计数器产生的顺序地址作为 **RAM** 的读出地址, 从交织地址生成器的输出的交织地址作为 **RAM** 的写入地址。

根据 802.16e 标准中的规定, 我们在发送端和接收端分别加入了 2 片自适应交织或去交织器来实现交织 1、交织 2 或去交织 1、去交织 2, 实现框图如图 8 所示。对于交织 1 采用 1 片自适应交织器, 状态转移如图 7 所示; 对于交织 2, 也采用 1 片自适应交织器 (当调制方式为 BPSK 和 QPSK 时此交织器不进行交织处理, 当调制方式为 16-QAM 和 64-QAM 时分别采用不同算法的交织, 交织地址算法如表 2 所示。交织与去交织器的区别仅仅在于用交织地址作为 **RAM** 的写入地址还是读出地址, 2 者总的结构相同 (如图 6), 只是选择的交织地址算法不同。为了加快交织器的处理速度, 还可以把交织地址生成器换成 ROM 来实现 (内部事先存入交织地址), 不过这种方案是以牺牲 FPGA 寄存器资源为代价的, 需要权衡利弊来决定是否采用这种办法。

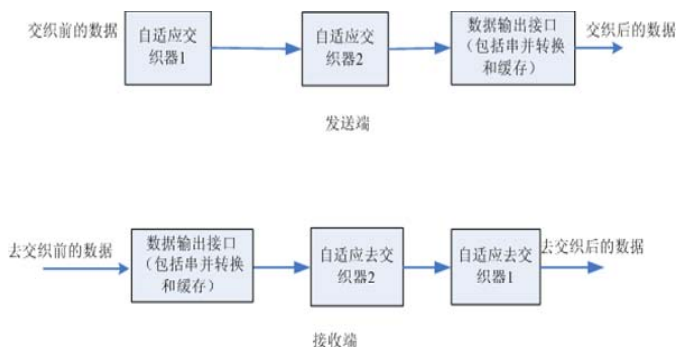


图 8 802.16e 交织部分实现总框图

4 设计验证与测试

本文用 VHDL 语言编程实现了基于表 1 的各种参数的自适应交织，并通过 Quartus7.2 软件仿真，链路联调验证。由于数据量大的缘故，仿真波形用图很难清楚表现出来，所以下面只列出仿真得到的部分数据结果：

(1) 交织参数为 16 个子信道、QPSK 调制：

交织前输入数据（16 进制）为：

3A 5E E7 AE 49 9E 6F 1C 6F C1 28 BC BD AB 57 CD BC CD E3 A7 92 CA 92 C2 4D BC
8D 78 32 FB BF DF 23 ED 8A 94 16 27 A5 65 CF 7D 16 7A 45 B8 09 CC

交织后输出数据（16 进制）为：

77 FA 4F 17 4E 3E E6 70 E8 CD 3F 76 90 C4 2C DB F9 B7 FB 43 6C F1 9A BD ED 0A 1C
D8 1B EC 9B 30 15 BADA 31 F5 50 49 7D 56 ED B4 88 CC 72 FC 5C

去交织后输出数据（16 进制）为：

3A 5E E7 AE 49 9E 6F 1C 6F C1 28 BC BD AB 57 CD BC CD E3 A7 92 CA 92 C2 4D BC
8D 78 32 FB BF DF 23 ED 8A 94 16 27 A5 65 CF 7D 16 7A 45 B8 09 CC

(2) 交织参数为 2 个子信道、16-QAM 调制：

交织前输入数据（16 进制）为：

EE C6 A1 CB 7E 04 73 6C BC 61 95 D3 B7 C4 EF 0E 4C 76 CF DC 70 69 B3 CE DB E0 E5
B7 B5 4E 88 7D A4 AE 31 30

交织后输出数据（16 进制）为：

B4 FF DA 06 E5 42 EC 1F 86 7C 29 93 9C AD 83 42 6B FE FC 6D CB F6 53 85 AE 68 22 7A
CE B1 E7 52 B0 EC BA 95

去交织后输出数据（16 进制）为：

EE C6 A1 CB 7E 04 73 6C BC 61 95 D3 B7 C4 EF 0E 4C 76 CF DC 70 69 B3 CE DB E0 E5
B7 B5 4E 88 7D A4 AE 31 30

(3) 交织参数为 1 个子信道、QPSK 调制：

交织前输入数据（16 进制）为：

EE C6 A1 CB 7E 04 73 6C BC 61 95 D3 B7 DF 00

交织后输出数据（16 进制）为：

BC EC A1 F4 8A 3A 7A 4F 78 39 53 87 DF 2A A2

去交织后输出数据（16 进制）为：

EE C6 A1 CB 7E 04 73 6C BC 61 95 D3 B7 DF 00

从以上 3 组数据中可以看到交织前的数据和数据自环回来去交织后的数据结果一致（数据流向如图 1 所示），从而验证了逻辑设计的正确性。

5 结论

分组交织在将某些突发错误随机化为随机错误的同时，也可能将某些周期性干扰变成突发错误，也可能因为交织不够充分使交织后仍然有一些相邻编码比特处于比较差的信道状态，

影响了译码性能。因此当编码在频域成帧时，需要采用自适应交织技术以提高系统性能。根据 802.16e 标准，本文给出了一种适用于 Wimax 系统的数据自适应交织的 FPGA 实现方案，并通过 Quartus7.2 软件仿真，链路联调验证，结果表明该方案是可行的。

参 考 文 献

- [1] IEEE Standard 802.16e™-2007. IEEE Standard for Local and Metropolilan Area Networks Part 16:Air Interface for Fixed Broadband Wireless Access Systems[S].2007
- [2] Altera Corporation. Cyclone III Device Handbook[K]. 2008.
- [3] Altera Corporation.Avalon memory-Mapped Interface Specification.v3.3[K]2007
- [4] Atmel Corporation. AT91RM9200 datasheet[K].2006
- [5] 张宗橙. 纠错码原理和应用[M]. 北京: 电子工业出版社, 2003.
- [6] 徐晓东, 井雅等.无线 OFDM 系统中自适应交织的设计[J].应用科学学报.2007, 25 (1): 103-107
- [7] 田耕, 徐文波等.无线通信 FPGA 设计[M].电子工业出版社.2008
- [8] 黄智伟.FPGA 系统设计与实践[M].北京: 电子工业出版社.2005:2,278-282

作者简介:

毛智强 (1983~), 男, 江苏南京人, 硕士, 重庆邮电大学研究生, 主要研究方向为个人通信。

雷维嘉 (1969~), 男, 云南元谋人, 副教授, 主要研究方向为无线通信技术。

谢显中 (1965~), 男, 四川人, 博士, 教授, 主要研究方向为移动通信技术、通信信号处理。

基于DTMB标准的卷积交织仿真与实现

肖剑辉 叶宇煌

(福州大学物理与信息工程学院 福建 福州 350002)

摘 要: 本文基于我国数字地面电视广播(DTMB)的传输标准,介绍了基于符号的时域卷积交织基本原理以及FPGA仿真与实现结果。

关键词: 数字电视; DTMB; 纠错编码; 交织; FPGA

Simulation and Implementation of Convolution Interleaving Based on DTMB

Xiao Jian Hui Ye Yu Huang

(College of Physic and Information Engineering, Fujian Fuzhou 350002)

Abstract: This Paper introduced the theory and method of implementation about symbol interleave in time domain, based on the Standard of DTMB.

Keywords: DTV; DTMB; FEC; Interleaving; FPGA

1 引言

2006年8月18日,随着《数字电视地面广播传输系统帧结构、信道编码和调制》(下文简称其为“国标”。)的发布,昭示了我国数字地面电视传输标准的确定。其中,该标准中所涉及到的时域交织方式,采用的是卷积交织法。

2 实现原理及要求

2.1 系统总图

如1图,为国标方案的发送端系统框图。在前向纠错之后,是星座映射与交织模块。使用交织的意义,就在于它把一段连续的突发差错离散成随机差错,方便前向纠错编码消除随机差错。

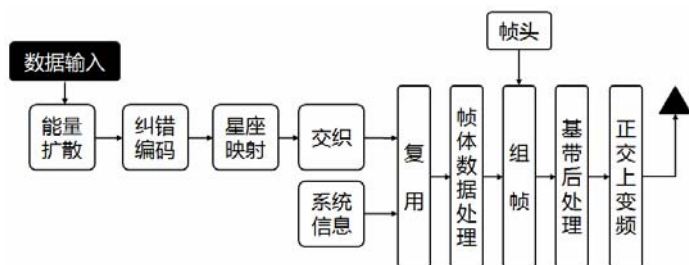


图 1

2.2 卷积交织原理

在国标中的时域交织，采用的是卷积交织方式，其原理图如下：

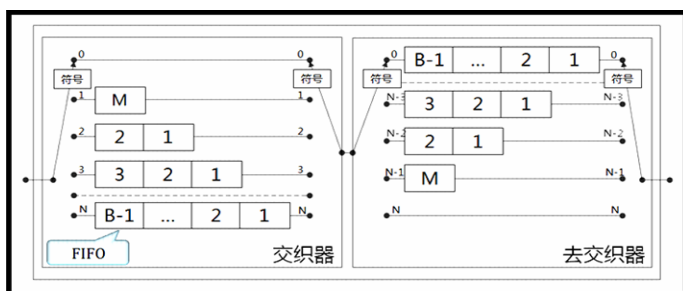


图 2

如上图所示，国标采用的卷积交织器，有 2 种工作模式，如图，B 为支路数，M 为每个移位寄存器的深度：

- a) 模式1：B=52，M=240符号，交织/解交织总延迟为170个信号帧。
- b) 模式2：B=52，M=720符号，交织/解交织总延迟为510个信号帧。

3 卷积交织的硬件实现

3.1 直接法

根据第 3 节的原理构架图，交织器事实上就是将数据采用不同大小的 FIFO 来进行输出延迟，故直接法即，采用 B 个不同深度的 FIFO 来实现。

该方法的交织器所占用的符号数为：

$$MEM = \frac{B \times (B-1) \times M}{2} \quad (1)$$

设每个符号为 T 比特。因此，直接法所需的寄存器数为：

$$MEM = \frac{B \times (B-1) \times M \times T}{2} \quad (2)$$

直接法的优点是，若交织深度较小，可以很简单直观地实现。然而对于 DTMB 标准中，B=52，M=720 或 240，如此大的交织深度下，此方法将占用大量的移位寄存器和众多的辅助电路，不便于实现。因此，放弃使用该方式。

3.2 计数器寻址法

从第 3 节的卷积交织的框图来看，采用的是二维寻址方式。

但从直接法的分析看出，在 DTMB 标准中，二维寻址方式不适用，因此，改而采用一维寻址方式，即：把 B 路的 FIFO 首尾相接起来，形成一个一维的 RAM。结构如图 3 所示：



图 3

B 路中的每一路第一个缓存位置占用一个相应基址，之后预留相应的空间作为该路缓存，缓存长度同 3 节所述。

在此方法中，原先的 FIFO 移位操作变成了对每一个单元的先写后读的操作，具体操作如下：

对于每一个将送入交织器缓存的数据，根据原理先运算出一维的 RAM 地址，再对其按先写后读的方式进行操作。

计数器寻址法的地址计算采用“基地址+偏移地址”。其中，第 k 个延迟支路的的基地址计算公式为：

$$f_k = \frac{k \times (k - 1) \times M}{2} + 1$$
$$k = 0, 1 \dots B - 1 \tag{3}$$

第 k 个延迟支路的偏移地址的计算公式为：

$$g_k = g_k + 1 \quad \text{当 } g_k \neq kM - 1 \text{ 时}$$
$$g_k = 0 \quad \text{当 } g_k = kM - 1 \text{ 时}$$
$$k = 0, 1 \dots B - 1 \tag{4}$$

故，地址计算公式为：

$$a_k = f_k + g_k$$
$$k = 0, 1 \dots B - 1 \tag{5}$$

如图 4，计数器寻址法的原理比较简单，在交织深度较小时，可以很方便地实现，然而在 DTMB 标准下的 M 取值 720 或 240，B 取值 52 情况下，将生成较多的计数器和辅助电路，占用大量硬件资源，降低工作效率，因此计数器寻址法不适用于此。

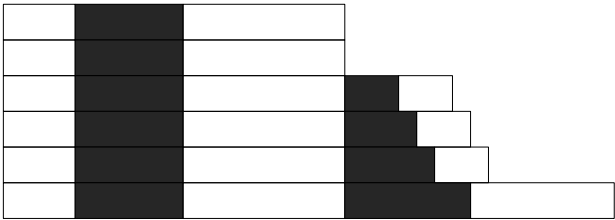


图 4

3.3 RAM矩形循环法

RAM 矩形循环法，即将输入数据先存入 RAM 中，不同行代表不同支路，通过控制读写来控制不同的延迟。而在 RAM 读写地址的控制上，有两种方法，第一种是将读写地址先存入一个 ROM 中，再根据 ROM 的输出决定 RAM 的读写地址，这种方法的优点在于寻址简单，缺点是需要占用一定的空间来存储地址；第二种是按照类似计数器寻址的方式，为了简便起见，在这里，取 $M=2$ ， $B=4$ 为例说明 RAM 矩形循环法的原理：

根据计数器寻址法，第 k 支路的总延迟为 kBM 个时隙，由此得出：

第 0 支路延迟为 0；第 1 支路延迟为 $BM=8$ ；第 2 支路为 $2BM=16$ ；第 3 支路为 $3BM=24$ ；如下图：

行	列	地址			
0	0	0	2	4	6
1	2	8	10	12	14
2	4	16	18	20	22
3	6	24	26	28	30

图 5

在进行交织时，对 RAM 的读写按照先写后读的方式进行，写地址按照序列：0，10，20，30，1，11，21，31……顺序写入，写满后返回首地址，如是循环；读地址按照 0，8，16，24，1，9，17，25……顺序读出，读到地址 31 后，返回 0 地址再进行下一周期的读出。

因此第 0 支路的数据延迟为 0，第 1 支路的数据延迟为 8，第 2 支路的数据延迟为 16，第 3 支路的数据延迟为 24。

在实现过程中，只需要利用两个计数器即可完成以上地址运算，分别为行计数器，和列计数器，设为 c ， r 。设写地址为 W ，读地址为 R 。

在每个数据到来时刻，进行写操作，行计数器 r 进行 $0 \sim B-1$ 的增 1 操作，至 $B-1$ ，转而进行读操作，同样是行计数器 r 进行 $0 \sim B-1$ 的加运算；以此作为第一个读写操作周期 T_k ，列计数器 c 以 T_k 作为自加周期，进行 $0 \sim M-1$ 的加法运算，当列计数器 c 计数至 $M-1$ 时，归 0，进行新的读写周期 $T = MT_k$ ：

$$r = r + 1 \tag{6}$$

$$c = c + 1 \quad (\text{以 } T_k \text{ 为运算周期}) \tag{7}$$

$$W = MB(r - 1) + M \times r + c \quad (\text{溢出本行时})$$
$$W = M \times B \times r + M \times r + c \quad (\text{未溢出本行}) \tag{8}$$

$$W = M \times B \times r + c \tag{9}$$

RAM 矩形循环法的优点在于：简单规则的寻址方式便于实现；缺点则在于占用了两倍于计数器寻址法的存储空间。

3.4 综合比较

上述 3 种卷积交织实现方式，在性能上作出如下比较：

方法	所需存储空间	控制电路	规模
直接法	$\frac{B \times (B - 1) \times M}{2}$	简单	很大
计数器寻址法	$\frac{B \times (B - 1) \times M}{2}$	复杂	大
RAM 矩形循环法	$B^2 \times M$	简单	小

图 6

综上所述，在设计中，虽然 RAM 矩形循环法占用了两倍的存储空间，但在当前片外存储器件价格相对低廉的情况下，可以作为首选方案考虑。

4 实现介绍

SDRAM 的全称是 Synchronous Dynamic Random Access Memory，意为同步动态随机存取存储器，SDRAM 从发展到现在已经经历了四代,从第一代的 SDR SDRAM 到第四代的 DDR3 SDRAM。

按照 4.1.3 节的方法，实现 DTMB 的卷积交织，所需的最大存储空间（符号）为：

$$B \times M \times B = 52 \times 720 \times 52 = 3182400$$

(10)

本设计选择 2M Words × 8 Bit × 4 Banks 的 IS42S16400 型(最高支持频率为 133MHz 的 SDR SDRAM 作为外部存储单元完成符号卷积交织。

IS42S16400 型 SDR SDRAM 的主要引脚及其功能如下：

名称	功能	名称	功能
CLK	Master Clock	DQM	DQ Mask Enable
CKE	Clock Enable	A0-11	Address Input
CS_N	Chip Select	BA0,1	Bank Address
RAS_N	Row Address Strobe	V _{DD}	Power Supply
CAS_N	Column Address Strobe	V _{DDQ}	Power Supply for DQ
WE_N	Write Enable	V _{SS}	Ground
DQ0-DQ15	Data I/O	V _{SSQ}	Ground for DQ

图 7

SDRAM 的优点是：1)价格低；2)体积小；3)速度快，4) 容量大，等。

缺点是：1) 掉电数据丢失；2) 需通过不断的电刷新来维持片内数据；3) 控制逻辑较为复杂。

由于 SDRAM 的控制复杂性,我们需引入 SDRAM 控制器方便用户对其进行操作。

4.1 SDRAM控制器

如图8，SDR SDRAM控制器的用户控制引脚(图4.6左侧)有：CLK，CMD[1:0]，CMDACK，ADDR，DATAIN，DM，DATAOUT。说明如下：

- CLK：SDRAM控制器的时钟输入；
- CMD：SDRAM控制器命令输入；
- CMDACK：SDRAM控制器命令响应；
- ADDR：SDRAM控制器地址输入；
- DATAIN：SDRAM控制器数据输入；
- DM：SDRAM控制器数据掩码输入；
- DATAOUT：SDRAM控制器数据输出。

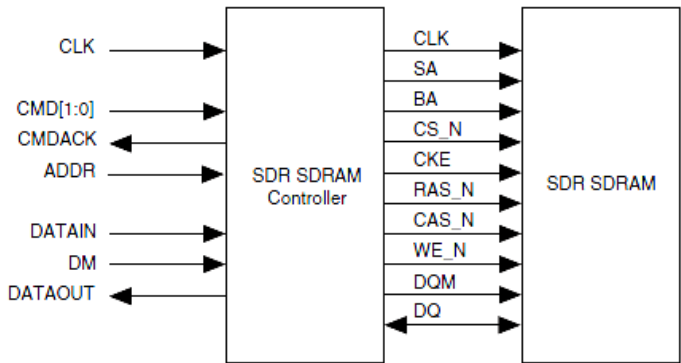


图 8

相应地，SDRAM 的控制引脚为图标 5.2 中 SDR SDRAM Controller 与 SDR SDRAM 相连部分，具体功能如上，其它关于 SDRAM 控制器的功能，此处不再赘述。

4.2 符号交织模块

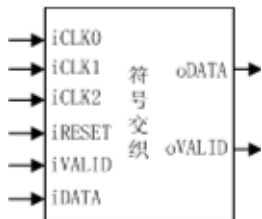


图 9

在FPGA设计上，采用该设计，如图该模块的的管脚说明如下：

iCLK0，数据写时钟输入； iCLK1，数据读时钟输入， iCLK2，SDRAM 刷新时钟输入； iRESET，异步清零输入； iVALID，数据有效信号输入； iDATA，交织符号数据输入； oVALID，数据有效信号输出； oDATA，交织符号数据输出。

按照4.3节的寻址方式，并通过SDRAM控制器进行对SDRAM的读写操作。数据先进入FIFO缓存，控制模块以FIFO半满信号作为触发控制交织进程开始信号，进行52个行读写操作

后,判断FIFO状态:当输入FIFO将空时,主控制模块则暂停交织进程,继而等待FIFO的半满信号再次到来,如此循环。

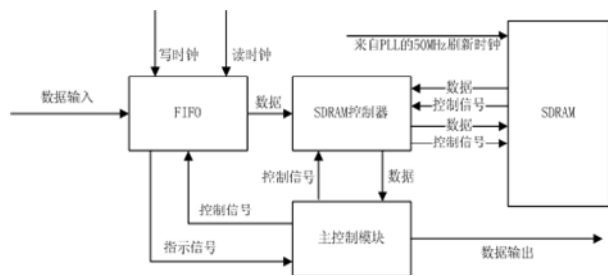


图 10

5 系统仿真

如图 11, 利用 DE2 研发板在 Quartus II 9.0 环境下的符号卷积交织 Signal Tap II 仿真图。

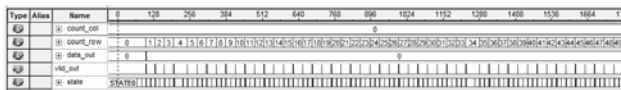


图 11

6 总结

通过以上的设计, 已完成了国标发送系统中的时域交织部分。本设计中的 SDRAM 的刷新频率定在了 50MHz。在针对 DTMB 调制器的全系统设计, 将整个国标发送系统的各个模块级联起时, 需更换 FPGA 芯片, 并采用更高速且廉价的 DDR2 SDRAM。

随着技术的发展, DTMB 很快将在我国各大城市普及, 我们应充分利用和完善数字电视地面国家标准, 推广更多的增值性应用, 为广大数字电视用户服务。

参 考 文 献

[1] 中国国家标准化管理委员会, GB20600-2006, 8-2006, “数字电视地面广播传输系统帧结构、信道编码和调制”, 中华人民共和国国家质量监督检验检疫总局, 2006

[2] 苏凯雄、郭里婷, 数字卫星电视接收技术, 北京: 人民邮电出版社, 2002

[3] 徐元欣、王匡、仇佩亮, 实现卷积交织的几种实用方法, 电路与系统学报, 第 6 卷第 1 期, 2001.3

[4] 段然、樊晓娅、张盛兵, 基于状态机的 SDRAM 控制器的设计和实现, 西北工业大学计算机学院, 2005

[5] 夏宇闻, Verilog 数字系统设计教程, 北京航空航天大学出版社, 2003

[6] 余涛, DVB-H 信道编码调制的设计与实现, 电子科技大学硕士学位论文, 2007

[7] 崔婷婷, 数字电视地面广播传输, DTMB 系统中交织技术的仿真与实现, 2008

[8] 蒲建发, 新型 DVB-C 信道编码、中频调制的全数字实现, 现代电子技术, 2006

[9] 张贇隆, 交织与解交织的算法研究及 FPGA 实现, 2007

[10] 潘松、黄继业, EDA 技术实用教程, 科学出版社, 2006

无线传感器网络中基于网络编码的信道 自适应协同传输方案

张 良 蔡跃明

(解放军理工大学通信工程学院, 江苏 南京 210007)

摘 要: 结合双向中继和物理层网络编码(PNC)技术, 提出一种信道自适应协同传输(CACT)方案。该方案根据节点之间信道状态信息从分布式空时编码(D-STBC)方案和协同辅助分集多址接入(C-NDMA)方案中进行选择, 获得了较小误帧性能。在给定的系统模型下, 分析了分别采用两种协同传输方案时误帧性能与中继节点位置的关系, 得出了在给定信噪比下为获得较小误帧性能而协同传输方案时的中继节点位置。

关键词: 无线传感器网络; 信道自适应协同传输; 物理层网络编码; 分布式空时码; 协同网络辅助分集

Network Coding Based Channel Adaptive Cooperative Transmission Scheme in Wireless Sensor Networks

Zhang Liang Cai Yue ming

(Institute of Communications Engineering, PLAUST, Nanjing 210007, China)

Abstract: In this paper, we propose a channel adaptive cooperative transmission (CACT) scheme combined with bi-directional relaying and physical layer network coding (PNC). This scheme can switch between PNC based distributed space-time block coding (D-STBC) scheme and cooperative network-assisted diversity multiple access (C-NDMA) scheme for the purpose of purchasing the minimized frame-error-ratio (FER) performance. Base on the analysis of the FER performance difference when the two schemes are adopted separately, we find the location of the relay where the system should change its cooperative scheme to obtain the smaller FER.

Keywords: WSN; CACT ; PNC; D-STBC; C-NDMA

1 引言

目前无线通信领域很多已有成果已经说明虚拟多入多出(V-MIMO)技术^{[1],[2]}通过联合单天线节点形成多天线阵列从而获得空间分集增益。但这要求信源与协同节点通过本地通信交

基金项目: 国家自然科学基金项目(No.60972051)、国家 863 科学基金项目(No.2009AA01Z249)和江苏省自然科学基金(BK2007002)资助课题。

互信息，如 Cui 等提出的分布式 Alamouti 空时编码方案^[3]需要两个时隙本地通信完成信息交互。而对于多源情况，Erick G 在[4]中设计了叠加编码调制策略，使两个信源节点在交叠发送的信号同时彼此互为协同。文献[5],[6]引入了专门用于协助信源交互信息的双向中继，进一步提高了分布式空时编码性能，但这些方案要求中继与信源之间具有良好信道条件来保证信息交互质量，而当这一条件不能满足时，若将 Lin 等在文献[7]提出的辅助分集多址接入(NDMA)思想移植到多源单中继协同传输模型中，利用中继重传信息在目的端进行信源冲突包解析，则同样能够获得空、时分集增益。

无线通信中的网络编码的核心思想是允许中继节点对经过它的数据进行解码和重编码，进一步提取或压缩信息后再转发出去以获得编码增益^[8]。物理层的网络编码（PNC）^[9]是一种处理电磁波信号接收和调制的物理层的网络编码。

本文将 PNC 应用到 D-STBC 与 C-NDMA 方案中提出一种信道自适应协同传输方案，此方案可根据节点之间 CSI 调整系统的协同策略，从而始终保持较低的接收误帧率。

2 系统模型

如图 1 所示，两个相邻簇头节点 A(B)在一个位于他们和目的端链路上的中继节点 R 协助下交互信息并与目的端 D 通信。为便于分析，将 A(B)到 D 距离归一化为 1，设 A(B)到 R 距离为 m ，R 到 D 距离为 $1-m$ 。考虑加性高斯白噪声（方差 $\sigma^2=1$ ）信道， h_{sr} 、 h_{sd} 、 h_{rs} 、 h_{rd} 、 h_{ds} 分别表征 A(B)到 R 和 D，R 到 A(B)和 D 以及 D 到 A(B)之间的信道系数，其中 $s \in \{a,b\}$ ，且服从大尺度衰落模型。因此 $h = d^{-\gamma} h_w$ ， d 为传输距离， γ 为衰落因子， h_w 服从 $CN(0,1)$ 分布。

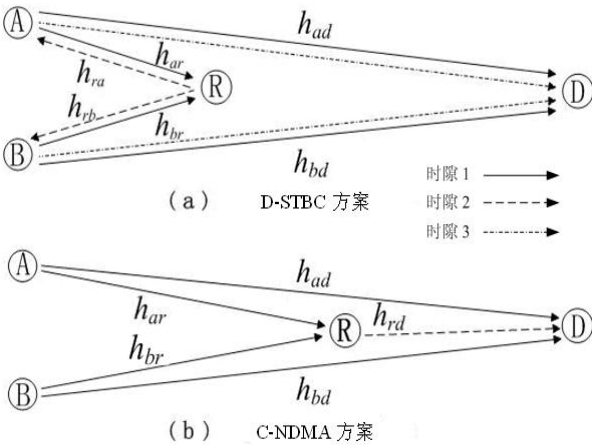


图 1 两种中继方案的系统模型

3 信道自适应协同传输方案

在给定的系统模型下，我们结合 PNC 和双向中继（Bi-directional relaying）^[10] 技术给出基于 PNC 的 D-STBC 和 C-NDMA 协同传输方案,并分析系统分别单独采用这两种方案时的误帧率性能。表 1 给出两种方案传输两个信源符号的时隙划分。本文用小写字母 x, y 表示信息

符号，对应的大写字母 X, Y 表示对应的等长度的信息帧。

表 1 D-STBC 与 C-NDMA 协同方案的时隙划分

方案	节点	时隙 1	时隙 2	时隙 3
D-STBC	A	发送 x_a	接收	发送 $-x_b^*$
	R	接收	发送 x_r	
	B	发送 x_b	接收	发送 x_a^*
C-NDMA	A	发送 x_a		
	R	接收	发送 x_r	
	B	发送 x_b		

3.1 基于PNC的D-STBC协同方案

本方案中, A(B)向 D 发送信息的同时利用 R 进行信息交互。完成信息交互后按照 D-STBC 方案将信息编码^[11]再发给 D, 这样接收端就可以获得空时编码带来的分集增益。整个通信过程可分为 3 个时隙 (见表 1)。

时隙 1, A(B)通过相互正交的信道同时向 D 发送第一个符号 $x_a(x_b)$, D 收到

$$y_{sd}^{(1)} = h_{ad}x_a + h_{bd}x_b + n_d \tag{1}$$

同时, 位于 A(B)到 D 路径上的 R 收到混合信号

$$y_{sr} = h_{ar}x_a + h_{br}x_b + n_{sr} \tag{2}$$

其中, n_d 和 n_{sr} 为噪声。

在第二时隙, R 采用 PNC 将收到的混合信号映射为 $x_r = x_a \oplus x_b$, 再向源与目的端广播, 而不是将混合信号直接放大转发^[12]或者分别解码出 x_a 与 x_b 的值再重新编码发送^[13]。

在考虑噪声的模型下, PNC 使用最小均方误差 (MMSE) 估计从混合信号中计算 $x_a \oplus x_b$ 的期望值

$$x_r = \alpha \mathcal{E} \{x_a \oplus x_b | y_{sr}\} = \alpha \sum_{t \in \{\pm 1 \pm j\}} t \Pr(t | y_{sr}) \tag{3}$$

式中, α 是对发送信号平均功率约束系数, $\mathcal{E} \{x_a \oplus x_b | y_{sr}\}$ 是接收信号为 y_{sr} 时 $x_a \oplus x_b$ 的 MMSE 估计,

$$\mathcal{E} \{x_a \oplus x_b | y_{sr}\} = \sum_{t \in \{\pm 1 \pm j\}} t \Pr(t | y_{sr}) = \sum_{t \in \{\pm 1 \pm j\}} \frac{t \Pr(y_{sr} | t) \Pr(t)}{\sum_{s \in \{\pm 1 \pm j\}} \Pr(y_{sr} | s)} = \frac{1}{4} \sum_{a,b,c,d \in \{\pm 1\}} \frac{(a \oplus c + j(b \oplus d)) P_{abcd}}{\sum_{a',b',c',d' \in \{\pm 1\}} P_{a'b'c'd'}} \tag{4}$$

其中 $j = \sqrt{-1}$, P_{abcd} 是当发送符号 $x_a = a + jb$ 且 $x_b = c + jd$ 时收到 y_{sr} 的先验概率, 由下式计算

$$P_{abcd} = \Pr(y_{sd} | x_a = a + jb, x_b = c + jd) = \varphi(y_{sd} - h_{ar}(a + jb) - h_{br}(c + jd), 1) \tag{5}$$

其中 $\varphi(\mu, \sigma^2) = e^{-|\mu|^2 / 2\sigma^2} / 2\pi\sigma^2$ 。

在考虑信道噪声的情况下, 由式(4)得到的 x_r 是 $x_a \oplus x_b$ 的估计值, 可以直接通过星座图映射得到判决符号。对一帧内所用符号通过(4)式进行估计, 可得到一帧的 MMSE 估计

$$\mathcal{E} \{X_a \oplus X_b | Y_{sr}\}, \text{平均功率为 } E\left(|\mathcal{E} \{X_a \oplus X_b | Y_{sr}\}|^2\right), \text{系数 } \alpha = \sqrt{1 / E\left(|\mathcal{E} \{X_a \oplus X_b | Y_{sr}\}|^2\right)}。$$

完成 PNC 编码以后, R 再向源节点广播 x_r , A(B)收到

$$y_{ra} = h_{ra}x_r + n_{rs}d \qquad y_{rb} = h_{rb}x_r + n_{rs} \tag{6}$$

在时隙 3，接收到 x_r 后，A(B)只需根据第一时隙发送的本地信号 x_{local} ，可计算另一个源节点发送信息符号

$$x_i = x_{\text{local}} \oplus x_r, \quad i \in \{a, b\} \quad (7)$$

当 A、B 都解调并译码出彼此在时隙 1 发送符号后，按照 2×1 V-MIMO 传输方案进行 Alamouti 编码^[6]并向 D 同时发送各自编码后的符号（见表 1），D 收到

$$y_{sd}^{(2)} = -h_{ad}x_b^* + h_{bd}x_a^* + n_d' \quad (8)$$

其中 n_d' 为复高斯噪声。

D 通过对两个时隙接收的信号 $y_{sd}^{(1)}$ 和 $y_{sd}^{(2)}$ 进行最大比合并^[14]（MRC）和极大似然检测（MLD）算法(9)检测出信源发送符号对 (a, b)

$$(a, b) = \arg \min_{(a, b) \in \{\pm 1 \pm j\}} (\|h_{ad}\|^2 + \|h_{bd}\|^2 - 1)(\|a\|^2 + \|b\|^2) + d^2(\tilde{x}_a, a) + d^2(\tilde{x}_b, b) \quad (9)$$

其中 \tilde{x}_a 、 \tilde{x}_b 为最大比合并结果

$$\tilde{x}_a = h_{ad}^* y_{sd}^{(1)} + h_{bd} y_{sd}^{(2)*}, \quad \tilde{x}_b = h_{bd}^* y_{sd}^{(1)} + h_{ad} y_{sd}^{(2)*} \quad (10)$$

$d^2(m, n)$ 为符号 m, n 之间的欧氏距离。

对于给定信道 h_{ad} 和 h_{bd} 而言， \tilde{x}_i ($i = a, b$) 仅仅是 x_i ($i = a, b$) 的函数，在给定每个节点发送功率相等的情况下，最大似然译码准则式 (9) 可简化为对于 x_i ($i = a, b$) 的两个独立的译码算法

$$a = \arg \min_{a \in \{\pm 1 \pm j\}} d^2(\tilde{x}_a, a), \quad b = \arg \min_{b \in \{\pm 1 \pm j\}} d^2(\tilde{x}_b, b) \quad (11)$$

3.2 基于PNC的C-NDMA协同方案

此方案第一时隙里仍由 A(B)同时向 D 发送各自信号，第二时隙中 R 将在上一时隙收到的混合信息进行 PNC 后重传给 D，而不是像传统 C-NDMA 方案那样直接转发给 D。由于 h_{ad} ， h_{bd} 和 h_{rd} 的非相关性，D 可直接利用连续两个时隙收到的信号进行冲突数据包联合解析。整个通信过程可分为 2 个时隙（见表 1）。

时隙 1，各节点的工作状态和信号模型同 D-STBC 方案，

$$y_{sd} = h_{ar}x_a + h_{br}x_b + n_d \quad (12)$$

时隙 2，中继 R 在 PNC 方案使用最大后验概率（MAP）从混合信号计算 $x_a \oplus x_b$ 的期望值

$$x_r = \beta \arg \max_{x_1 \oplus x_2 \in \{\pm 1 \pm j\}} \Pr(y_3 | x_1 \oplus x_2) P(x_1 \oplus x_2) \quad (13)$$

等同于 D-STBC 方案， β 为 x_3 的平均功率约束系数， $P(x_1 \oplus x_2)$ 是 $x_1 \oplus x_2$ 的先验概率分布，QPSK 调制下， $P(x_1 \oplus x_2) = 1$ ，则可推得

$$x_r = \arg \max_{a, b, c, d \in \{\pm 1\}} \{P_{a'b'c'd'} + P_{(-a')b'(-c')d'} + P_{a'(-b')c'(-d')} + P_{(-a')(-b')(-c')(-d')}\} \quad (14)$$

其中 P_{abcd} 见(5)式。

完成 PNC 映射得到 x_r 以后，不同于 D-STBC 方案的是 R 直接向 D 广播而不是 A(B)，D 收到

$$y_{rd} = h_{rd}x_r + n_{rd} = h_{rd}x_a + h_{rd}x_b + n_{rd} \quad (15)$$

根据连续两个时刻收到的分别来自 A(B)和 R 的信号，D 可以用 MRC 和 Soft MAP 检测信息符号 x_a ， x_b 。以下分析以符号 x_a 的检测为例（ x_b 可根据对称性类比得到），MRC 后相位最大化的信号为

$$y_a = \theta_{ad}^* y_{sd} + \theta_{rd}^* y_{rd} \quad (16)$$

其中 $\theta_{rd} = h_{rd} / \|h_{rd}\|$, $\theta_{ad} = h_{ad} / \|h_{ad}\|$ 。Soft MAP 检测时使用[4]中定义的对数似然比 (LLR) 分别对 x_a , x_b 进行符号判决,

$$L(x_a | y_a) = \log \left(\frac{\sum_{x_b \in \{1, -1\} | x_a = 1} \exp \left(-\frac{1}{2\sigma^2} (y_a - \theta_{ad}^* h_{bd} x_b - \theta_{rd}^* h_{rd} x_b - \theta_{ad}^* h_{ad} x_a - \theta_{rd}^* h_{rd} x_a)^2 \right) P(x_a) P(x_b)}{\sum_{x_b \in \{1, -1\} | x_a = -1} \exp \left(-\frac{1}{2\sigma^2} (y_a - \theta_{ad}^* h_{bd} x_b - \theta_{rd}^* h_{rd} x_b - \theta_{ad}^* h_{ad} x_a - \theta_{rd}^* h_{rd} x_a)^2 \right) P(x_a) P(x_b)} \right) \quad (17)$$

如果 $L(x_a | y_a) > 0$, 判定 $x_a = 1$, 否则 $x_a = -1$ 。式中 σ^2 为噪声方差, $P(x_a)$, $P(x_b)$ 分别是符号 x_a , x_b 是先验概率, QPSK 调制下 $P(x_a) = P(x_b) = 1/4$ 。虽然 $P(x_a)P(x_b)$ 可在(17)中约分, 但这里为表达不失一般性仍然将其包括。

3.3 基于PNC的CACT方案

我们将基于 PNC 的 C-NDMA 协同方案命名为方案 1, 而基于 PNC 的 C-NDMA 协同方案命名为方案 2。通过前文的分析可知, 方案 1 利用分布式空时分组码获得空时分集增益, 在信源信息交互的高效性和准确性能够保证的前提下, 本文第三部分的仿真说明其误帧性能优于方案 2。而文献[15]的结果说明, 当中继到目的端 CSI 良好时方案 2 能获得较高增益。

据此, 我们设计自适应协同方案, 当信源 A、B 同时有信息需要发送时, 首先在尽可能靠近自身的范围内寻找能量状况和信道条件较理想的节点作为中继 R。如果能够找到, 则系统执行方案 1, 否则扩大搜寻范围继续寻找中继。直到找到的中继 R 到 D 的信道衰落更小于 R 到 A(B)的, 系统应及时做出决策执行方案 2 以获得较优性能。

接下来的章节中, 我们在大尺度衰落模型下考虑中继的位置对节点之间信道衰落系数的影响, 通过对两种协同策略误帧率性能相对于中继位置的仿真, 寻找系统应做方案改变时的中继位置。

4 仿真结果及讨论

仿真采用 QPSK 调制, 每个固定长度数据帧包含 1024bits 信息。衰落因子 $\gamma = 3$ 。根据 A、B 相对于 R 和 D 具有对称性的假设, 有 $|h_{ar}|^2 = |h_{br}|^2$, $|h_{ra}|^2 = |h_{rb}|^2$, $|h_{ad}|^2 = |h_{bd}|^2$ 。

图 2 给出了方案 1 和方案 2 在中继 R 距离 A(B) m 分别为 0.3 和 0.7 时的误帧性能。从曲线可知, 当 R 靠近 A(B)时, 由于 A(B)到 R 信道衰落较小, 方案 1 误帧性能优于方案 2。在误帧率为 10^{-2} 处性能增益达到 5dB, 因此本方案中当信源传输信息时首先在附近寻找信道条件较好的节点以执行方案 1。但如果 R 远离 A(B)而接近 D, 则方案 2 更可靠, 在 10^{-1} 处误帧性能超过方案 1 约 2.5dB。

为了直观观察中继位置对于误帧性能影响, 图 3 给出误帧率性能相对中继位置变化曲线, 信噪比分别取 10dB 和 15dB。由结果可知: 对于方案 1 来说, R 到 A(B)的距离越近, 传输越可靠, 而对于方案 2 来说正好相反; 方案 1 平均误帧性能优于方案 2; 两方案性能曲线在 $m=0.7$ 左右相交。应以此为参考点做出自适应方案选择, 找更可靠的方案进行协同传输。不同的信噪比下, 两方案性能曲线交点有所不同。也就是说, 本方案的理论上的自适应选择参考点根据当前系统信噪比会有所不同。

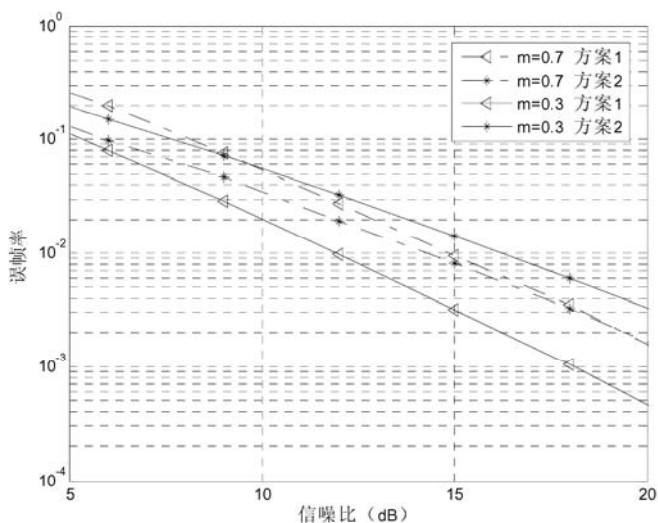


图2 中继分别靠近源和目的端时两方案误帧率比较

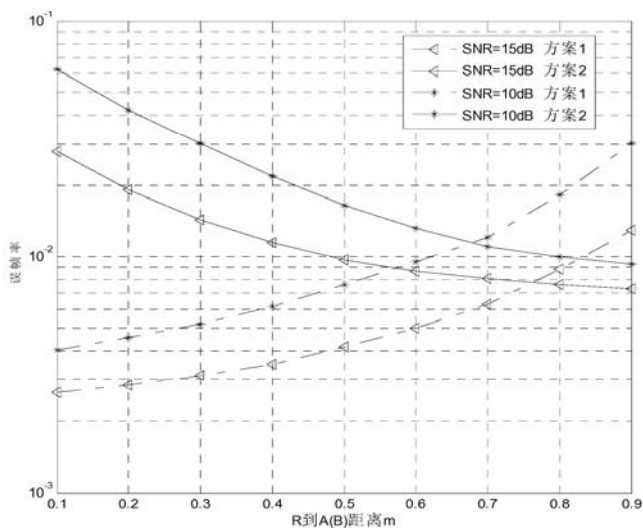


图3 误帧率曲线中继位置的关系

5 结论

本文设计了一种 WSN 中基于 PNC 的信道自适应协同传输方案, 该方案能根据信源、中继、目的端之间的信道衰落情况自适应选择 D-STBC 或 C-NDMA 协同传输方案。根据对两种协同方案的分析, 我们找到两者在中继位置不同时的性能差异, 并使系统在合适的时机做出决策以获得更优的误帧率性能。本文的结论对于 WSN 中的协同策略选择问题有实际意义。下一步工作方向是在方案自适应选择时, 综合考虑系统复杂度、速率以及能效等因素, 根据当前网络业务需求做决策, 而不是仅仅以稳定性为唯一追求目标。

参 考 文 献

- [1] S. K. Jayaweera, Virtual MIMO-based Cooperative Communication for Energy-constrained Wireless Sensor Networks [J], *IEEE Trans on Wireless Communication*, vol. 5, no. 5, pp. 984-989, May. 2006.
- [2] S. Alamouti, A simple transmit diversity technique for wireless communications [J], *IEEE J. Sel. Areas Commun.*, pp. 1451-1458, Oct. 1998.
- [3] S. Cui, A. J. Goldsmith and A. Bahai, Energy-efficiency of MIMO and cooperative MIMO techniques in sensor networks [J], *IEEE Journ. Select. Areas. Commun.* vol 22, pp.1089-1098, Aug. 2004.
- [4] Erik G. Larsson and Branimir R. Vojcic, Cooperative Transmit Diversity Based on Superposition Modulation [J], *IEEE Commun Letters*, vol. 9, no. 9, Sep. 2005.
- [5] Ling Lv, Hongyi Yu and Jianzu Yang, Opportunistic Cooperative Network-Coding Based on Space-Time Coding for Bi-Directional Traffic Flows [A], *IEEE Nuclear Science Record, Hong Kong, 2006*.
- [6] J. Nicholas Laneman and Gregory W. Wornell, Distributed Space-Time-Coded Protocols for Exploiting Cooperative Diversity in Wireless Networks [J], *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2415-2425, Oct. 2003.
- [7] Lin R, Petropulu A P, A New Wireless Network Medium Access Protocol Based on Cooperation [J] *IEEE Trans on Signal Process*, vol. 53, no. 12, pp. 4675-4684, Dec. 2005.
- [8] S.-Y. R. Li, R.W. Yeung, and N. Cai, Linear network coding [J], *IEEE Trans. on Inform. Theory*, vol. 49(2), pp. 371-381, 2003.
- [9] S. Zhang, S. C. Liew, and P. P. Lam. Hot topic: physical layer network coding [A], in *Proc. 12th MobiCom*, pp. 358-365, New York, NY, USA, 2006.
- [10] P. Popovski and H. Yomo, Bi-directional amplification of throughput in a wireless multi-hop network [A], in *Proc. IEEE VTC'06*, Oct. 2006.
- [11] L. Zhang, Chengkang Pan, and Yueming Cai, An Opportunistic Cooperative Multi-Source Transmission Scheme in Wireless Sensor Networks [A], *IEEE International Conference on Wireless Communications and Signal Processing, Nanjing, China. Dec 2009*.
- [12] P. Popovski and H. Yomo, Wireless Network Coding by Amplify-and-Forward for Bi-Directional Traffic Flows [J], *IEEE Commun Letters*, vol. 11, no. 1, Jan. 2007.
- [13] S. Zhang, S. -C. Liew, Joint Design of Physical-layer Network Coding and Channel coding [R], Technique Report On line: <http://arxiv.org/ftp/arxiv/papers/0807/0807.4770.pdf>.
- [14] Branka Bucetic, Jinhong Yuan, *Space-Time Coding* [M], Hoboken NJ: John Wiley and Sons, 2003.
- [15] Yang H, Petropulu A P. ALLIANCES with optimal relay selection [A], *IEEE International Conference on Acoustics Speech and Signal Processing*, Honolulu, Hawaii, pp.49-52, April 2007.

作者简介:

张 良 (1984-), 男, 安徽合肥人, 硕士研究生, 研究方向为无线传感器网络;

蔡跃明 (1961-), 男, 福建永泰人, 解放军理工大学通信工程学院, 博士生导师, 研究方向为下一代移动通信, 无线资源管理等。

双向中继对秩亏MIMO信道容量的影响

屠 佳¹ 蔡跃明^{1, 2}

(1. 解放军理工大学通信工程学院, 江苏 南京 210007;

2. 东南大学移动通信国家重点实验室, 江苏 南京 210096)

摘 要: 在视距 (LOS, line-of-sight) 环境中, MIMO 信道矩阵的秩亏特性严重抑制了其信道容量的增加。本文针对秩亏 MIMO 信道, 使用基于放大转发协议的双向中继增加了 MIMO 信道矩阵的秩, 同时, 由于在两个节点之间双向同时传输数据, 提高了频谱利用率。理论分析和仿真表明, 与不使用中继辅助的 MIMO 信道相比, 所提出方案在 LOS 环境中能够有效提高秩亏 MIMO 信道的容量, 并且其所需的时隙仅为传统放大转发中继的一半。

关键词: 多输入多输出 (MIMO, Multiple Input Multiple Output); 秩亏信道; 双向中继; 信道容量

Impact of Bi-directional Relays on the Capacity of Rank-Deficient MIMO Channels

Tu Jia¹ Cai Yue ming^{1, 2}

(1. Institute of Communications Engineering, PLA University of Science and Technology, Nanjing 210007, China; 2. Nation Mobile Communications Research Laboratory, Southeast University, Nanjing 210096, China)

Abstract: In a line-of-sight environment the MIMO channel matrix is rank-deficient and therefore the capacity increase diminishes. Aiming at the rank-deficient MIMO channels, the amplify-and-forward bi-directional relay is used to increase the rank of the MIMO channels and effectively reduce the loss in spectral efficiency by bi-directional simultaneous transmission of data between the two nodes. In a line-of-sight environment, we investigate and compare the performance of a bi-directional relay assisted MIMO channel with a MIMO channel without relay nodes. Simulation results show that using bi-directional relays can solve the problem of rank-deficient MIMO channels and it requires half of the number of the time slots compared with the conventional amplify-and-forward relays.

Keywords: MIMO; rank-deficient channel; bi-directional relay; capacity of channel

1 引言

在无线通信系统中, 多入多出 (Multiple Input Multiple Output, MIMO) 技术能够在不增加带宽的情况下大幅度提高系统容量, 给无线通信带来新的发展空间^{[1],[2]}。未来无线通信系统需要支持更高数据速率传输, 工作频带也将超过 5GHz, 这就造成 MIMO 技术在实际应用中将面

临其瓶颈问题的挑战：对富散射环境的需求。MIMO 信道容量的增加是依赖于传播环境的，当信道系数相关性逐渐增大时，仅有几条子信道对整个系统总容量做出贡献，系统性能将大大降低。然而在高频环境下，传播信道中多为视距(line-of-sight, LOS)传输，信道相关性不断增加，很难进行非视距传输，特别是在纯 LOS 环境下，不存在多径传输，MIMO 信道矩阵的秩仅为 1，这使得我们利用空时复用技术的可能性为零。为了解决这个问题，很多学者进行了深入研究，从不同角度利用不同方法来改进信道矩阵的秩，从而提高信道的传输速率。文献[3],[4]从发射端进行考虑，通过预先设计某种特定的阵列天线来最大化信道矩阵的秩。而从信道的传输过程来看，中继的引入成为了一种有效的解决方法。从本世纪七十年代开始，Van der Meulen 在文献[5]、Sato 在文献[6]以及 Cover 和 Gamal 在文献[7]中相继对中继信道的容量问题进行了讨论，但是这些讨论都不是针对 LOS 信道进行的。文献[8]提出了一种新的思想：将协同中继随机分布，充当传输信道中的散射体，形成富散射环境，从而有效解决了 LOS MIMO 信道矩阵的秩亏问题。类似的，文献[9]在 Rayleigh 信道下讨论了 MIMO 中继对信道容量的影响，并分析了这种 MIMO 中继信道的一种极限性能（中继数量趋于无穷多）。

到目前为止，已经出现了很多利用协同中继解决 MIMO 信道秩亏问题的方法，但是我们发现，绝大部分的研究都是基于传统的半双工中继方案。在这种方案下，若想在节点 A 和节点 B 之间交换信息，需要四个时隙才能完成，换句话说，传统的半双工中继方案在频谱效率上有所损失。对此，文献[10]提出了一种双向中继方案：在第一个时隙，节点 A 和节点 B 同时发送数据到中继节点，此时中继接收到的是节点 A 和 B 的叠加信息；在第二个时隙，中继将这个叠加信息放大转发给节点 A 和 B 进行接收。对于节点 A(或 B)来说，它可以通过从接收到的叠加信息中去除自干扰来检测出节点 B(或 A)发送的信息，从而完成节点间的信息交互。所以说，相比于传统的半双工中继，双向中继方案仅需要其一半的时隙。文献[11]即讨论了高斯 MIMO 双向信道的容量问题，但这个讨论的前提是富散射环境。

我们的贡献：本文主要是针对放大转发双向中继对秩亏 MIMO 信道容量的影响进行了讨论。利用文献[8]中将传统的协同中继作为散射体的思想，我们将双向中继引入秩亏 MIMO 信道，形成富散射环境，从而有效提高信道矩阵的秩。理论分析和仿真表明，与不使用中继辅助的 MIMO 信道相比，该方法在 LOS 环境中能够有效提高秩亏 MIMO 信道的容量，同时，由于使用了双向中继，完成两个节点间的信息交互仅需要传统中继一半的时隙，有效提高了频谱利用率。

本文后续安排如下：第 2 节给出了基于放大转发协议的双向中继方案。这节是本文的重点，我们首先介绍了系统模型，然后具体讨论秩亏 MIMO 信道的容量问题；第 3 节给出了计算机仿真结果，比较分析算法的各种性能；最后在第 4 节中进行了小结。

符号表达： $\mathbf{0}$ 和 \mathbf{I} 分别表示零矩阵和单位阵， $(\mathbf{a})^T$ 和 $(\mathbf{a})^H$ 分别表示转置矢量（或矩阵）和共轭转置矢量（或矩阵）， $\|\mathbf{a}\|$ 表示 Frobenius 范数。

2 双向中继方案

正如前文所述，在 LOS 环境中，LOS 分量严重影响了整个信道响应矩阵的秩。因此在本节中，我们引入双向中继来解决信道秩亏的问题。图 1 给出了一个协同双向中继网络图。图中，节点 A 和 B 分别有 N_A 和 N_B 根天线。在这个系统中，两个节点通过中继辅助利用两个时

隙进行信息交换。在第一个时隙，节点 A 和节点 B 同时发送自己的数据到中继节点，此时中继接收到的是节点 A 和 B 的叠加信息；在第二个时隙，中继将这个叠加信息放大转发，节点 A 和 B 分别进行接收、检测出对方节点发送的信息。在图 1 中我们可以看到，双向中继是随机分布的，因此这些中继可以被看成一组“主动的”全向散射体，为系统形成一种多径信道。这里，之所以我们将中继称为“主动”散射体，是由于它在起到普通散射体作用的同时，也将噪声加入了转发信号中。

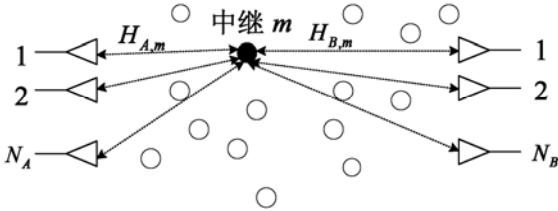


图 1 协同双向中继网络

2.1 系统模型

如图 1 所示，节点 A 和 B 通过一个中继节点建立一条双向连接。节点 A 和 B 分别配有 N_A 和 N_B 根天线，每个中继配有 N_R 根天线，整个系统中一共使用 M 个中继来辅助通信，从而提高 MIMO 信道矩阵的秩。然而，实际使用中继节点并不是必须配备多根天线的，为了分析简单，在我们的仿真中设定 $N_R=1$ 。假设节点 A 和 B 之间的直传信道为 LOS 信道。

在时隙 k ，节点 A 和 B 同时传输数据给中继节点。这里，我们将节点 A 和 B 的传输信号分别表示为 $\mathbf{x}_{AB}^{(k)}=[x_{AB,1}^{(k)},\dots,x_{AB,N_A}^{(k)}]^T$ 和 $\mathbf{x}_{BA}^{(k)}=[x_{BA,1}^{(k)},\dots,x_{BA,N_B}^{(k)}]^T$ ，二者均服从功率约束，即 $E\{(\mathbf{x}_{AB})^H(\mathbf{x}_{AB})\}=P_A$ ， $E\{(\mathbf{x}_{BA})^H(\mathbf{x}_{BA})\}=P_B$ 。在中继 m ($m=1,\dots,M$) 处的接收信号 $\mathbf{r}_{R,m}^{(k)}$ 可以表示为

$$\mathbf{r}_{R,m}^{(k)} = \mathbf{H}_{A,m} \mathbf{x}_{AB}^{(k)} + \mathbf{H}_{B,m} \mathbf{x}_{BA}^{(k)} + \mathbf{n}_{R,m}^{(k)} \quad (1)$$

其中， $\mathbf{H}_{A,m} \in \mathbb{C}^{N_R \times N_A}$ 和 $\mathbf{H}_{B,m} \in \mathbb{C}^{N_R \times N_B}$ 分别表示中继 m 和节点 A 与 B 之间的信道状态系数。中继 m 上的噪声为高斯白噪声，其分布为 $\mathbf{n}_{R,m} \sim \mathcal{CN}(0, \sigma_R^2 \mathbf{I}_{N_R})$ 。在下一时隙传输之前，预先将接收矢量 $\mathbf{r}_{R,m}^{(k)}$ 乘以一个增益矩阵 \mathbf{G}_m 。为了简化表示，我们定义下面一些矢量和矩阵：

$$\mathbf{r}_R^{(k)} = [\mathbf{r}_{R,1}^{(k)T}, \dots, \mathbf{r}_{R,M}^{(k)T}]^T \quad (2)$$

$$\mathbf{n}_R^{(k)} = [\mathbf{n}_{R,1}^{(k)T}, \dots, \mathbf{n}_{R,M}^{(k)T}]^T \quad (3)$$

$$\mathbf{G} = \text{diag}(\mathbf{G}_1, \dots, \mathbf{G}_M) \quad (4)$$

$$\mathbf{H}_A = [\mathbf{H}_{A,1}^T, \dots, \mathbf{H}_{A,M}^T]^T \quad (5)$$

$$\mathbf{H}_B = [\mathbf{H}_{B,1}^T, \dots, \mathbf{H}_{B,M}^T]^T \quad (6)$$

因此，在时隙 k ，在中继以及节点 A 和 B 上接收到的信号分别表示为：

$$\mathbf{r}_R^{(k)} = \mathbf{H}_A \mathbf{x}_{AB}^{(k)} + \mathbf{H}_B \mathbf{x}_{BA}^{(k)} + \mathbf{n}_R^{(k)} \quad (7)$$

$$\mathbf{r}_A^{(k)} = \mathbf{r}_B^{(k)} = \mathbf{0} \quad (8)$$

这里需注意的是，在时隙 k 节点 A 和 B 不能接收任何信号，因为从天线的实际角度考虑，此时 A 和 B 正在发送信号，因此在同一物理信道上，这些天线不能再进行任何其他接收操作。

在下一时隙 $k+1$ ，中继节点将经过放大的矢量 $\mathbf{G}\mathbf{r}_R^{(k)}$ 分别发送给节点 A 和 B。不失一般性，我们这里只考虑节点 A 上的处理问题，节点 B 上具有相同的操作（这里不再赘述）。一般来说，对于 \mathbf{G}_m 的选择主要有两种，分别表示如下：

$$\mathbf{G}_m = \sqrt{\frac{P/M}{\|\mathbf{H}_{A,m}\|^2 P_A/N_A + \sigma_R^2}} \quad (9)$$

$$\mathbf{G}_m = \sqrt{\frac{P/M}{\sigma_{A,m}^2 P_A/N_A + \sigma_R^2}} \quad (10)$$

其中， $P = P_A + P_B$ ， P/M 表示中继 m 上的最大传输功率。通常我们将式(9)称为可变增益，而式(10)则为固定增益，因为在式(9)中 $\mathbf{H}_{A,m}$ 是瞬时信道状态信息值，它是在不断变化之中的，而式(10)中 $\sigma_{A,m}^2$ 在一定时间内基本保持不变，从而有利于整个系统进行长期的功率约束控制。因此，在本文中，我们选择式(10)作为中继节点的增益矩阵。在时隙 $k+1$ ，在节点A上的接收信号可以表示为：

$$\mathbf{r}_A^{(k+1)} = \mathbf{H}_A \mathbf{G} \mathbf{r}_R^{(k)} + \mathbf{n}_A^{(k+1)} \quad (11)$$

将式(7)代入式(11)，并整理可得：

$$\mathbf{r}_A^{(k+1)} = [\mathbf{H}_A \mathbf{G} \mathbf{H}_A \quad \mathbf{H}_A \mathbf{G} \mathbf{H}_B] \begin{bmatrix} \mathbf{x}_{AB}^{(k)} \\ \mathbf{x}_{BA}^{(k)} \end{bmatrix} + [\mathbf{H}_A \mathbf{G} \quad \mathbf{I}] \begin{bmatrix} \mathbf{n}_R^{(k)} \\ \mathbf{n}_A^{(k+1)} \end{bmatrix} \quad (12)$$

2.2 秩亏MIMO信道的容量

图 2(a)是对式(12)的一个图解说明。从式(12)中我们可以看到，由于增益矩阵 \mathbf{G} 和信道矩阵 \mathbf{H}_A 的存在，在时隙 $k+1$ ，接收端上的噪声不再是高斯白噪声，而变成有色噪声。如图 2(a)中所示，令 $\mathbf{n}_c^{(k+1)} = \mathbf{H}_A \mathbf{G} \mathbf{n}_R^{(k)} + \mathbf{n}_A^{(k+1)}$ 表示有色噪声分量。 $\mathbf{n}_c^{(k+1)}$ 的自相关矩阵 Λ 可以表示成：

$$\Lambda = E[\mathbf{n}_c^{(k+1)}(\mathbf{n}_c^{(k+1)})^H] = \sigma^2 \left(\mathbf{H}_A \mathbf{G} \mathbf{G}^H \mathbf{H}_A^H \frac{\sigma_r^2}{\sigma^2} + \mathbf{I}_{N_A} \right) = \sigma^2 \Lambda' \quad (13)$$

类似于文献[8]中给出的具体分析，我们可以得到一个白化矩阵：

$$\mathbf{W} = \mathbf{D}^{-\frac{1}{2}} \mathbf{V}^H \quad (14)$$

其中， $\Lambda' = \mathbf{V} \mathbf{D} \mathbf{V}^H$ ， $\mathbf{V} \mathbf{V}^H = \mathbf{V}^H \mathbf{V} = \mathbf{I}_{N_A}$ 。

图 2(b)描述了一个从图 2(a)转化而来的等效信道模型，其白噪声表示为 $\mathbf{w}^{(k+1)} = \mathbf{W} \mathbf{n}_c^{(k+1)}$ ，分布为 $\mathbf{w}^{(k+1)} \sim CN(0, \sigma_A^2 \mathbf{I}_{N_A})$ 。令 $\hat{\mathbf{r}}_A^{(k+1)}$ 表示等效接收信号，则可以表示为：

$$\begin{aligned} \hat{\mathbf{r}}_A^{(k+1)} &= \mathbf{W} \mathbf{r}_A^{(k+1)} \\ &= [\mathbf{W} \mathbf{H}_A \mathbf{G} \mathbf{H}_A \quad \mathbf{W} \mathbf{H}_A \mathbf{G} \mathbf{H}_B] \begin{bmatrix} \mathbf{x}_{AB}^{(k)} \\ \mathbf{x}_{BA}^{(k)} \end{bmatrix} + \mathbf{w}^{(k+1)} \end{aligned} \quad (15)$$

令 $\hat{\mathbf{H}} = \mathbf{W} \mathbf{H}_A \mathbf{G} \mathbf{H}_A$ ， $\tilde{\mathbf{H}} = \mathbf{W} \mathbf{H}_A \mathbf{G} \mathbf{H}_B$ ，并将两个时隙联合考虑，可以得到如下的矢量组：

$$\begin{bmatrix} \mathbf{r}_A^{(k)} \\ \hat{\mathbf{r}}_A^{(k+1)} \end{bmatrix} = \underbrace{\begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \hat{\mathbf{H}} & \tilde{\mathbf{H}} \end{bmatrix}}_{\mathbf{H}_1} \begin{bmatrix} \mathbf{x}_{AB}^{(k)} \\ \mathbf{x}_{BA}^{(k)} \end{bmatrix} + \begin{bmatrix} \mathbf{0} \\ \mathbf{w}^{(k+1)} \end{bmatrix} \quad (16)$$

从式(16)中我们可以发现，若要对双向信道进行容量计算，将会遇到这样一个问题：对于节点A来说， $\hat{\mathbf{H}}$ 实际是一个干扰，它不仅不会为MIMO系统的容量做出任何贡献，反而会降低

节点A处的信道容量。因此为了能够有效增加复合（两个时隙同时考虑）信道矩阵的秩，形成特征值分布，我们需要将信道矩阵 \mathbf{H}_1 转变为另一个形式。将式(16)改写成式(17)后，干扰项 $\hat{\mathbf{H}}$ 将不再对特征值分布产生影响。

$$\begin{bmatrix} \mathbf{r}_A^{(k)} \\ \hat{\mathbf{r}}_A^{(k+1)} \end{bmatrix} = \underbrace{\begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \tilde{\mathbf{H}} & \hat{\mathbf{H}} \end{bmatrix}}_{\mathbf{H}_{eq,1}} \begin{bmatrix} \mathbf{x}_{BA}^{(k)} \\ \mathbf{x}_{AB}^{(k)} \end{bmatrix} + \begin{bmatrix} \mathbf{0} \\ \mathbf{w}^{(k+1)} \end{bmatrix} \quad (17)$$

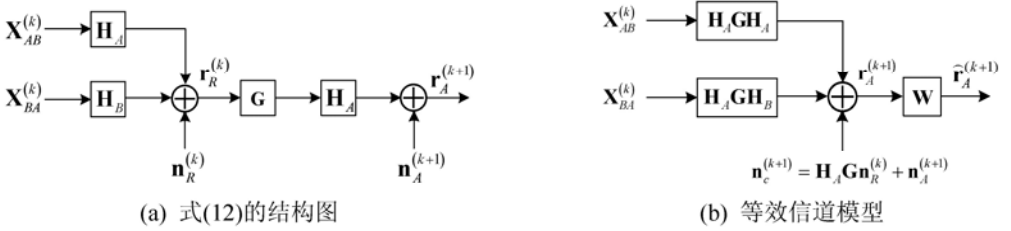


图 2 在时隙 $k+1$ ，节点 A 上接收信号的结构图

信道矩阵 $\mathbf{H}_{eq,1}$ 描述了一个具有加性高斯白噪声的 $2N_A \times (N_A + N_B)$ 的 MIMO 信道。尽管 \mathbf{H}_A 的秩仅为 1，但通过上述变换，复合信道矩阵 $\mathbf{H}_{eq,1}$ 的秩显然得到了增加。因此，由文献[1]可得，在节点 A 上复合信道的容量为

$$C_{H,A} = \log_2 \det \left(\mathbf{I}_{2N_A} + \frac{P}{N_A \sigma_A^2} \mathbf{H}_{eq,1} \mathbf{H}_{eq,1}^H \right) \quad (18)$$

同样，在节点 B 也有类似的操作。因此节点 B 上复合信道的容量可以表示为

$$C_{H,B} = \log_2 \det \left(\mathbf{I}_{2N_B} + \frac{P}{N_B \sigma_B^2} \mathbf{H}_{eq,2} \mathbf{H}_{eq,2}^H \right) \quad (19)$$

其中， $\mathbf{H}_{eq,2} = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{W} \mathbf{H}_B \mathbf{G} \mathbf{H}_A & \mathbf{W} \mathbf{H}_B \mathbf{G} \mathbf{H}_B \end{bmatrix}$ 。

3 仿真分析

3.1 仿真条件

本文针对双向中继系统在秩亏 MIMO 信道条件下的性能进行仿真验证。其中节点 A 和 B 分别位于 $(-1,0)$ 和 $(1,0)$ 。同时，多个中继在由两节点作为焦点、长半轴为 4、短半轴为 2 的椭圆内随机分布。节点 A 和 B 处有多个天线，为简化分析，我们假设双向中继处只有一根天线。与文献[8]相同，本文采用单径衰落信道模型，其功率路径衰落指数为 $\alpha = 4$ 。

3.2 性能分析

本文采用双向中继以实现节点 A、B 之间的信息交互，因此，节点 A 和 B 在信息交互的两个时隙里不直接互传信号。这就是说，我们忽略节点 A、B 之间的直传链路，但是为了验证双向中继在 LOS MIMO 信道条件下的性能，这里假设：在上述两个时隙中节点 A 和 B 仍直接交互信息。首先，节点 A 在第 k 个时隙通过直传链路向节点 B 发送信号，然后节点 B 在第

$k+1$ 个时隙向节点 A 发送信号，我们称之为“直传”。

如图 3 所示，我们在平均接收信噪比（SNR）相同的情况下比较双向中继辅助不同种类 MIMO 链路的信道容量，这里令 $\text{SNR}=20$ 。第一类为瑞利衰落 MIMO 信道，我们采用双向中继辅助节点 A 和 B 之间的信息交互，并利用该类性能仿真曲线作为容量的上边界。第二类为 LOS MIMO 信道，我们仍然采用双向中继方案，并选择两种不同的中继数目，这也是本文的工作所在。第三类为无中继辅助“直传”信道，我们可以将该类性能曲线视为容量的下边界。由图 3 可知，在 LOS MIMO 信道条件下，随着节点 A 和 B 天线数目的增加，本文所提出的中继方案与“直传”方案相比，具有更大的容量和更好的系统性能。但是，从图中我们可以发现本文所提出的方案在 LOS MIMO 信道条件下的性能劣于瑞利衰落 MIMO 信道，这是因为本文选择 $M=2N_A$ 和 $M=4N_A$ ，当 N_A 很小， M 没有足够大到可以提供丰富的散射环境而造成的。然而，通过比较本文提出方案的两条性能曲线可知，随着中继数量的增加，其作为“主动”散射体用来构成一种多径信道的作用将越来越明显。

系统误码率性能如图 4 所示。在节点 A 上检测出节点 B 发送信号的过程为：首先，节点 A 从接收信号中去除自己发送信号的干扰，然后利用 MMSE 频域均衡器来重构节点 B 的信号。为了尽可能从理论上提高 LOS MIMO 信道的性能，与文献[8]相同，这里我们取 $N_A=N_B=8$ ， $M=60$ 。然而，在实际中为了避免使用过多的中继，我们可以利用门限准则[8]来避免使用那些性能较差、无用的中继。由图 4 可知，当使用直传方案时，系统误码率性能很差，这是由于 LOS MIMO 信道矩阵的秩为 1，直传方案无法利用 MIMO 信道的容量，因此该方案的性能从本质上来说不能得到改善。然而，本文所提出方案的误码率性能却远远优于“直传”方案，且与瑞利 MIMO 信道相比仅有 2dB 的损失。这就说明使用双向中继作为主动“散射体”可以提高 LOS MIMO 信道的秩，增加对信道容量的利用能力。

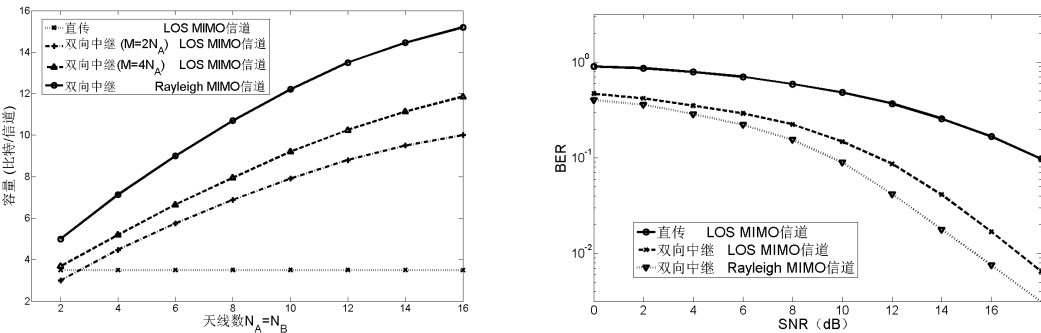


图 3 双向中继辅助不同 MIMO 信道的容量比较 图 4 不同传输方案用于不同 MIMO 信道的 BER 性能比较

以上分析说明本文所提出的方案对于秩亏 MIMO 信道来说，可以充分利用 LOS MIMO 信道的容量。但是，由于节点 A 接收的是叠加后的信息，因此与传统中继方案相比，在检测信息时的解码复杂度会有一定的增加。近期，很多研究成果针对双向中继的解码问题给出了各自的解决办法，如文献[12]、[13]。这里，我们不再赘述，仅认为通过较为简单的检测过程节点 A 可以准确的检测到节点 B 发送的信息，而实际上，我们也能够找到合适的译码器，将其应用到本文所提出方案。

4 结论

本文主要研究了多个放大转发双向中继对秩亏 MIMO 信道容量的影响问题。考虑到在一个 LOS 环境中, MIMO 信道矩阵呈现严重的秩亏特性, 导致信道容量大大降低, 为了解决这个问题, 我们使用了协同双向中继来增加 MIMO 信道矩阵的秩。由于使用了双向中继辅助两个节点之间同时进行数据传输, 仅需要两个时隙即可完成节点之间的信息交互, 使得我们的方法较之传统的半双工中继(需要四个时隙)而言, 切实提高了频谱利用率。仿真结果表明, 使用双向中继能够解决 MIMO 信道的秩亏问题, 有效提高 LOS MIMO 信道的容量。

参 考 文 献

- [1] E. Telatar, "Capacity of multi-antenna Gaussian channels," *Eur. Trans. Telecommun.*, vol. 10, pp. 585-595, Nov. 1999.
- [2] G. J. Foschini, "Layered space time architecture for wireless communication in a fading environment when using multi-element antennas," *Bell Labs. Tech. J.*, vol. 1, pp. 41-59, 1996.
- [3] I. Sarris and A. R. Nix, "Design and performance assessment of high-capacity MIMO Architectures in the Presence of a Line-of-Sight Component," *IEEE Trans. Vehicular Technology*, vol. 56, pp. 2184-2202, July 2007.
- [4] F. Bohagen, P. Orten and G. Oien, "Construction and capacity analysis of high-rank line-of-sight MIMO channels," in *IEEE Wireless Communications and Networking Conference (WCNC)*, Mar. 2005.
- [5] E. van der Meulen, "Three-terminal Communication Channels," *Adv. Appl. Prob.*, vol. 3, pp. 120-154, 1971.
- [6] H. Sato, "Information Transmission through a Channel with Relay," *The Aloha System*, University of Hawaii, Honolulu. Tech. Rep. B76-7.
- [7] T. M. Cover and A. El Gamal, "Capacity theorems for the relay channel," *IEEE Trans. Inform. Theory*, vol. 25, pp. 572-584, Sept. 1979.
- [8] A. Wittneben and B. Rankov, "Impact of Cooperative Relays on the Capacity of Rank-Deficient MIMO Channels," in *Proc. 12th IST Summit on Mob. and Wirel. Comm.*, pp. 421-425, Jun. 15-18, 2003.
- [9] H. Bolcskei, R. Nabar, O. Oyman and A. Paulraj, "Capacity scaling laws in MIMO relay networks," *IEEE Trans. Information Theory*, vol. 5, pp. 1433-1444, June 2006.
- [10] Y. C. Liang and R. Zhang, "Optimal analogue relaying with multi-antennas for physical layer network coding," in *Proc. IEEE ICC 2008*, pp. 3893-3897, May 2008.
- [11] R. F. Wyrembelski, T. J. Oechtering, I. Bjelakovic, C. Schnurr and H. Boche, "Capacity of Gaussian MIMO Bidirectional Broadcast Channels," in *Proc. IEEE International Symposium on Information theory (ISIT 2008)*, Canada, July 2008.
- [12] T. H. Pham, Y. C. Liang, A. Nallanathan and G. H. Krishna, "On the Design of Optimal Training Sequence for Bi-Directional Relay Networks," in *IEEE Signal Processing Letters*, vol. 16, No.3, pp. 200-203, March 2009.
- [13] T. Unger and A. Klein, "Linear transceiver filters for relay stations with multiple antennas in the two-way relay channel," in *Proc. 16th IST Mobile and Wireless Summit*, Budapest, Hungary, Jul. 2007.

基于ENC28J60 芯片的以太网接口方案设计

付衍荣¹ 房颖²

(1. 美国微芯商贸(上海)有限公司厦门分公司,福建 厦门 361001;

2.福州大学物理与信息工程学院, 福建 福州 350108)

摘要: 随着网络和嵌入式系统的发展, 嵌入式以太网技术在人们的工作生活中已经起到越来越重要的作用。介绍了以实现智能家庭报警系统为背景, 采用 Microchip 16 位单片机与以太网控制器 ENC28J60 芯片实现以太网接口电路的设计与编程方法。这种方法实现的以太网通信接口硬件电路简单、可靠性高、成本低, 具有很好的实用价值。

关键词: 嵌入式系统; 以太网接口; 安防监控报警; ENC28J60

Design of Ethernet Interface Based on ENC28J60

Fu Yan rong¹ Fang Ying²

(1. Microchip Technology Trading(Shanghai)Co.,Ltd Xiamen Branch, Xiamen ,Fujian, 361001;

2.School of Physics and Information Engineering, Fuzhou University, Fuzhou, Fujian, 350108)

Abstract :With the continuous development of the network communication technology and embedded system, Embedded Ethernet technology has been played an increasingly important role in people's working life .In this paper, in order to achieve a smart home security alarm monitoring system, the design of Embedded Ethernet interface based on the Microchip 16bit MCU and the ENC28J60 is described. This Ethernet communication interface has the characteristics of simple hardware circuit, high reliability , low cost . It is widely utilized.

Keywords: Embedded system; Ethernet interface; Security alarm monitoring; ENC28J60

1 引言

随着以太网的迅猛发展, 各种家电设备、仪器设备正在逐步走向网络化,以实现网络信息资源的共享, 进行远程控制等。因此, 在数据采集和控制设备中采用 TCP/IP 网络协议标准, 实现基于网络的远程数据采集、控制、故障诊断、上传/下载数据等功能具有重要的实际意义。本文以实现智能家庭报警系统为背景, 提出了基于 ENC28J60 芯片的以太网接口方案设计。

智能家庭报警系统(如图 1)主要包括有由探测器报警输入、控制单元、以太网通信、报警输出、32 字符液晶屏、4 个发光二极管、EEPROM 等组成, 而各种安防探测器通过有线或无线接口与单片机控制器相连, 当检测到有非法人员入侵、发生火灾等时, 就会立即向单片机控制器发出报警信号。单片机控制器会判断何处发生报警, 并立即现场报警, 且将报警信息通过以太网传到电信服务器, 服务器端收到信息后可选择以短信或电话报警形式通知业主或

公安人员。本文主要介绍采用 PIC24FJ256GA106 单片机,运用以太网控制芯片——ENC28J60,实现以太网接口电路的设计与编程方法。这种方法与传统上使用专为计算机系统设计的以太网控制器实现以太网通信的方法相比,采用的是单 MCU 的方式,硬件接口电路相对简单,体积小,可靠性高,而且成本低。

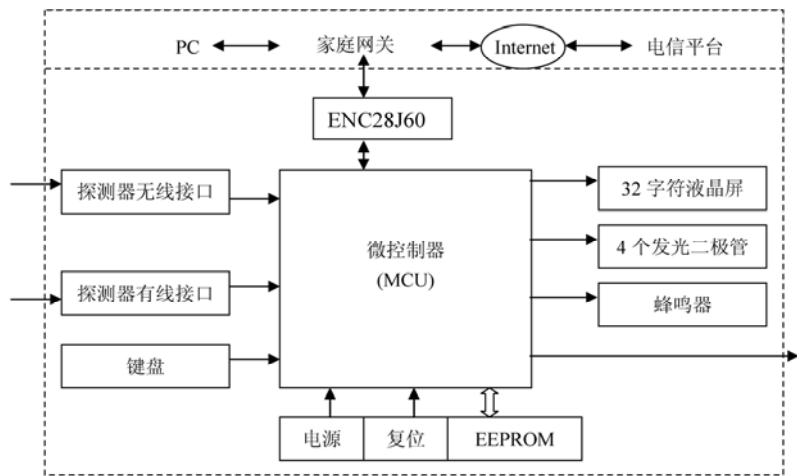


图 1 报警系统结构示意图

2 硬件电路设计

2.1 主要器件选择

微控制器选用的是高性能 PIC 单片机 PIC24FJ256GA106, PIC24FJ256GA106 是 Microchip 公司开发的 16 位通用闪存单片机,具有成本低,设计灵活,可靠性高等特点。PIC24FJ256 MCU 模块采用 16 位(数据)改良的哈佛架构,并带有增强型指令集,工作速度达到 16MIPS,满足运行多种网络协议的要求,而且具有丰富的外设接口,其 SPI 接口允许与外设进行高速的同步数据传输,这些外设器件可以是串行 EEPROM,移位寄存器,显示驱动器等^[1]。

以太网控制芯片选用 ENC28J60。目前在嵌入式系统的开发中,比较常用的独立以太网控制器采用的封装均超过 80 脚,如 RTL8019AS、CS8900A 均为 100 脚,这些器件不仅结构复杂、体积大,而且价格相对昂贵,系统开销也较大。而 Microchip 推出 28 引脚独立以太网控制器 ENC28J60,可为嵌入式系统提供低引脚数、低成本、精简的远程通讯解决方案。图 2 为典型的 ENC28J60 接口电路。

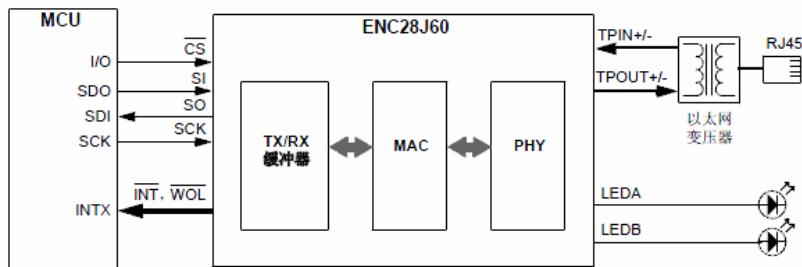


图 2 典型的 ENC28J60 接口电路

初始化后进入主程序循环,主要包括单片机的控制程序和网络数据传输程序。而网络数据传输主要包括对要发送的数据按照以太网数据帧格式进行封装并发送以及对接收的数据帧进行解包两个方面。

3.2 发送数据包

ENC28J60 内的 MAC 在发送时会自动生成前导符和帧起始定界符。此外,MAC 可根据配置生成填充(如果需要)和 CRC 字段。主控制器必须生成所有其他帧字段,并将它们写入缓冲存储器,以待发送。此外,ENC28J60 还要求在待发送的数据包前添加一个包控制字节。

要进行数据包发送时,主控制器应该正确编程 ETXST 指针,使之指向存储器中未用的单元;并且写入包控制字节、目标地址、源 MAC 地址、类型/长度和数据有效负载;还需要正确编写 ETXND 指针,指向数据有效负载的最后一个字节。发送数据包的流程图如图 5 所示。

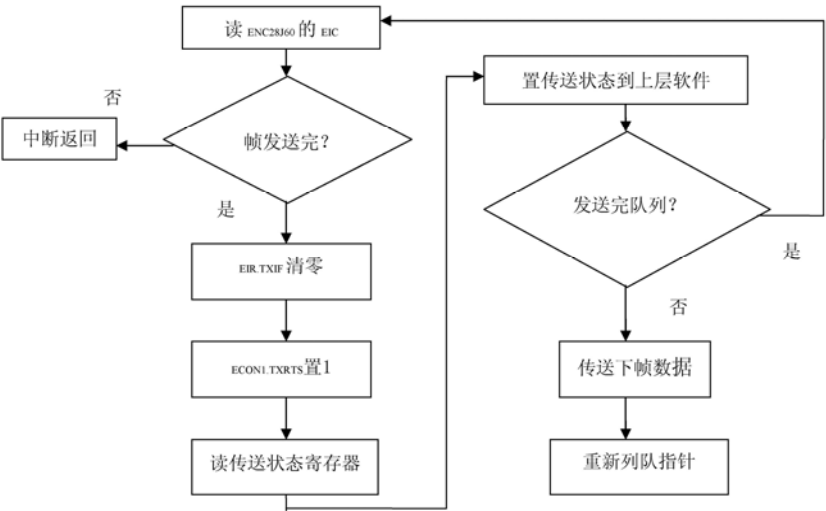


图 5 发送数据包流程图

3.3 接收数据包

假设接收缓冲器已完成初始化,MAC 已正确配置而且接收过滤器已配置为接收以太网数据包后,主控制器如果需要在接收到数据包时产生一个中断,就要将 EIE.PKTIE 和 EIE.INTIE 置 1;如果需要在由于缓冲空间不足导致数据包丢失时产生一个中断,就要将 EIR.RXERIF 清零,并将 EIE.RX-ERIE 和 EIE.INTIE 置 1;并且要将 ECON1.RXEN 置 1 使能接收。在将 RXEN 置 1 后,将不能修改双工模式和接收缓冲器起始和结束指针。此外,要阻止不期望接收的数据包,在更改接收过滤器配置寄存器(ERXFCON)和 MAC 地址前将 RXEN 清零。接收数据包的流程如图 6 所示。

通过对 ENC28J60 的软件驱动正确设置后,嵌入 TCP/IP 协议,就可以实现了以太网的接入,再辅以相应的应用层协议就可以实现远程控制功能。本系统已实现了 HTTP 协议、FTP 协议等。本系统采用静态网页访问的形式,使用 HTML 语言来编写网页,运用 Microchip 提供的 MPFS.exe 工具上传网页,网页代码存储在 EEPROM 中。嵌入式系统把要传输的数据、参数等存入 EEPROM 中,当有客户端请求连接时,这些数据信息通过 HTTP 协议沿 TCP 连接发送至浏

览器,这样就可以在远端 PC 机的浏览器中访问存储在系统中的网页,从而实现远程访问及控制,如图 7 所示。

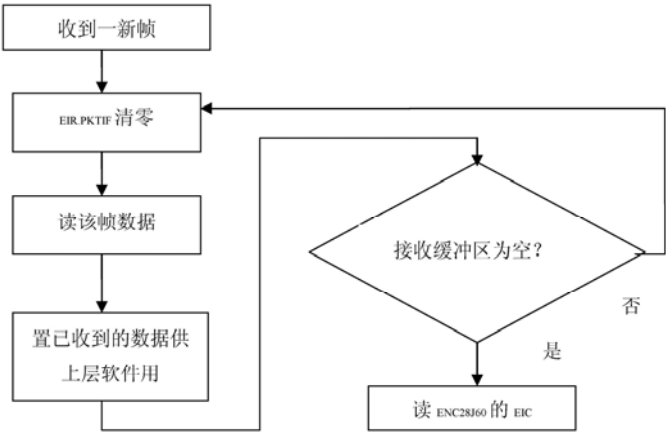


图 6 接收数据包流程图



图 7 存储于系统中的网页

4 结束语

本文介绍了一种基于 ENC28J60 芯片的以太网接口方案设计,用它可以方便地实现嵌入式系统的以太网连接。智能家庭报警系统就是以此为基础,加入与实际需求相关的协议,按需要进行配置完成系统功能,实现了基于电信宽带网络平台进行报警及全面的联动处警,且在测试中性能稳定。

随着数字化、智能化仪器的飞速发展,采用以太网进行通信的应用将会越来越广泛。利用 ENC28J60 芯片将网络技术与单片机技术的完美结合,可广泛适用于智能交通、智能家居、工业控制、信息家电、医学仪器等各种嵌入式系统应用场合。

参 考 文 献

- [1] PIC24FJ256GA110 系列数据手册 具有外设引脚选择功能的 64/80/100 引脚 16 位通用闪存单片机 [EB/OL]. <http://www.microchip.com/>.2008 年.
- [2] ENC28J60 数据手册 带 SPI 接口的独立以太网控制器[EB/OL]. <http://www.microchip.com/>.2006 年.
- [3] 郁继宗,彭树生.基于 PIC 单片机的以太网数据采集与控制电路设计[J].微计算机信息(测控自动化).2005 年第 21 卷第 2 期

作者简介:

付衍荣(1975-),男,福建省浦城县,学士,美国微芯商贸(上海)有限公司厦门分公司,资深应用工程师。

手机: 013809505118 email: fred.fu@microchip.com

地址: 厦门市厦禾路 189 号银行中心 2608 室 邮编: 361001

房颖(1981-),女,福建省福州市,硕士,福州大学物理与信息工程学院,助教,计算机通信。

手机: 013003886688 email: fangying@tom.com

地址: 福州市闽侯上街大学城福州大学物理与信息工程学院通信工程系 邮编: 350108

基于多频带OFDM的矿井巷道超宽带无线通信建模

钱建生 王 鑫 孙彦景 李世银

(中国矿业大学 信息与电气工程学院 江苏 徐州 221008)

摘 要: 针对无线通信在煤矿复杂环境条件下由多径传播所造成的频率选择性衰落的问题, 提出基于多频带正交频分复用 (MB-OFDM) 构建适合矿井复杂环境的超宽带无线通信系统。为适应煤矿巷道, 采用独立的角度脉冲响应建立基于 S-V 模型的信道模型。结合修正的信道模型, 对基于 MB-OFDM 的矿井 UWB 系统进行建模, 给出了频带的划分方式、信道的改造办法以及无线系统的设计等。仿真分析表明: 基于多频带 OFDM 的超宽带无线通信能有效抵抗矿井巷道多径衰落的影响, 具有抗噪声能力强, 误码率低, 覆盖范围大的优势。

关键词: 超宽带; 信道模型; 煤矿巷道; 多频带; 角度脉冲响应

The Model and Simulation of UWB Based on MB-OFDM in Coal Mine Underground

Qian Jian sheng Wang Xin Sun Yan jing Li Shi yin

(School of Information & Electronic Engineering, China University of Mining and Technology, XuZhou, Jiangsu 221008,China)

Abstract: For the frequency selective fading of UWB caused by the spread of multi-path in coal mine environment, UWB based on MB-OFDM in coal mine underground was proposed. In order to make up for the traditional S-V model's defects in fading channel, the traditional channel impulse response was changed into independent angle impulse response, and the revised channel model was built. With the revised channel model, a model of the UWB based on MB-OFDM was built, and gave the band division scheme, the transformation approach for channel, and the program of wireless system. The model of the system was simulated in the complex environment of coal mine. The results show: the system can be effective against multipath fading, and has high anti-noise ability, low bit-error rate, and long coverage.

Keywords: UWB; channel model; tunnel of coal mine; multiple band; angle impulse response;

0 引言

超宽带无线通信 (UWB) 系统具有传输速率高、低功耗、抗干扰等优点, 在煤矿巷道复杂环境下具有很好的应用前景^[1-5], 文献[1,2]对矿井环境下, 无线信号的传播特性进行研究。文献[3]对超宽带无线系统在巷道通信中的应用进行分析, 从理论上说明 UWB 系统在矿井巷道中的应用。文献[4]从信道路径损耗的角度分析了 UWB 信号在矿井巷道中的传播特性。这

些研究中,并没有给出具体的系统模型。且各个 UWB 设备之间或 UWB 设备与其它无线通信设备之间的干扰情况不同,以及不同巷道的干扰情况也不同,仍然是 UWB 系统在矿井中应用所面临的问题。

为了对矿井 UWB 系统性能进行精确评估,必须要建立能够真实反映在多径传播环境下 UWB 信号传播特性并且又准确的 UWB 通信系统。为了有效克服矿井巷道的路径损耗、多径性能、信道能量衰减不同等特点^[5]。本文基于多频带正交频分复用(MB-OFDM)技术抗多径时延的特性,结合修正的信道损耗模型,建立了基于 MB-OFDM 的矿井 UWB 系统,在矿井复杂环境下,能够同时保证超宽带速率高,且满足抗干扰的要求。

1 信道模型

在煤矿井下通信环境中,接收信号由经过反射、绕射和散射的多径信号叠加而成,本文选用 IEEE802.15.3a 推荐的基于簇的 S-V 模型作为信道模型^[5,6]。

针对煤矿无须认证的无线环境,为了避免复杂性、降低功耗。本文采用基于时频交织(TFI)的 MB-OFDM-UWB 的第 1 子频带组,即第 1、2、3 子频带,中心频率分别为 3432MHz,3960MHz 以及 4488MHz^[7,8],主要使用 3.168GHz~4.752GHz 前三个子带。考虑到多径信道是典型的频率选择性信道,不同的频段和带宽会有不同的频率响应特性和时间分辨率,有必要对信道加以改造,使改造后的信道既体现不同子带的特性,同时又适应不同的信息速率。要达到既满足抽样定理又最大限度地减小数据量的目的,满足数字系统的传输时钟 T_p ,要进行信道插值和抽取。可根据频率成分对不同的频段采用相应的抽样率。

设频率特性为 $h_1(t)$,其多径分辨率仍为 0.167ns,要将其变成分辨率为 T_p 的 $h'(t)$,也就是将抽样率转变为 T_p/T 倍,先对信号作 N 倍的插值,经过低通滤波器,再作 M 倍的抽取, $T_p/T=N/M$,过程如图 1 所示。



图 1 信道改造过程

2 系统建模

2.1 修正的信道模型

具体来讲, S-V 模型的信道冲击响应可以表示为^[7]:

$$h(t)=X\sum_{n=1}^N\sum_{k=1}^{K(n)}\alpha_{nk}\delta(t-T_n-\tau_{nk}) \tag{1}$$

其中 X 是对数正态随机变量,代表信道的幅度增益; N 是观测到的簇的数目, $K(n)$ 是第 n 簇内接收到的多径数目, α_{nk} 是第 n 簇中第 k 条路径的系数, T_n 是第 n 簇到达时间; τ_{nk} 是第

基金项目: 国家自然科学基金项目(50904070); 江苏省高新技术重大项目(BG2007012); 中国矿业大学科研基金项目(OC080303).

n 簇中第 k 条路径的时延。

在时域范围内室内多径传播信道，表现为成簇分布的规律，且跟多径分量的到达时间和到达角度密切相关。因此信道模型写成跟时间及方位角有关的脉冲响应表达式：

$$h(t,\theta)=X\sum_{n=1}^N\sum_{k=1}^{K(n)}\alpha_{nk}e^{j\theta_{nk}}\delta(t-T_n-\tau_{nk}) \tag{2}$$

如果多径分量的到达时间和到达角度之间存在相关性，那么簇均方的更大角度的偏移将导致更长的时延。而国外的研究实验中发现其信道特性并不满足这一点，因此在这里可以认为时间和角度之间是独立的^[4]，因此可以将 $h(t,\theta)\approx h(t)h(\theta)$ ，将 $h(\theta)$ 完全独立出来，故：

$$h(\theta)=\sum_{n=1}^N\sum_{k=1}^{K(n)}\delta(\theta-\omega_{nk}) \tag{3}$$

ω_{nk} 是第 k 簇中第 n 径的到达角度。 $h(\theta)$ 是一个独立的角度脉冲响应，可以认为该响应是一个与信道的时延脉冲响应类似的模型。

表 1 给出了矿井巷道内探测信道的环境特征和仿真所需的信道模型参数。主要径的取法通常是取功率在峰值功率(最强径)的 10dB 内，或功率和为总功率的 85%的径。因此为了确定参数值，首先必须求出这些特征参数的统计量。得到这些统计量之后，再通过拟合的方法即可得上述参数值。由于矿井巷道与室内各种信道的环境差异较大，经过拟合得到的参数也反映出不同特点。

表 1 探测信道的环境特征和信道模型参数

信道数量	信道特征	Λ (1/ns)	λ (1/ns)	Γ
500	极端恶劣	0.0667	2.1	36
Υ	σ_{ϵ} (dB)	σ_{ϵ} (dB)	σ_g (dB)	类型
24	3.3941	3.3941	3	LOS

采用上述参数，对该信道环境下信道进行仿真，得到的冲激响应，如图 2 所示：

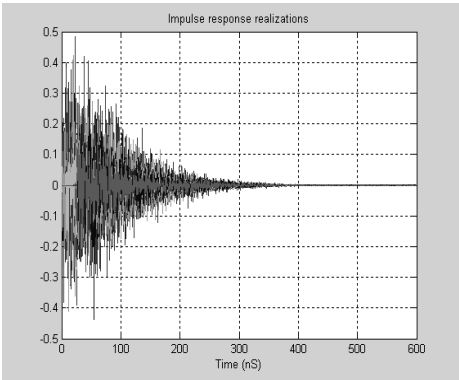


图 2 离散信道冲激响应

在图中能明显观察到多径成簇，这与实际测量中 UWB 多径成簇现象相验证。这是多径衰落信道中的典型结果，可以看到接收功率呈现指数衰减。在 10ns 后，功率衰减趋于零,在短时间内达到减小功耗的效果。

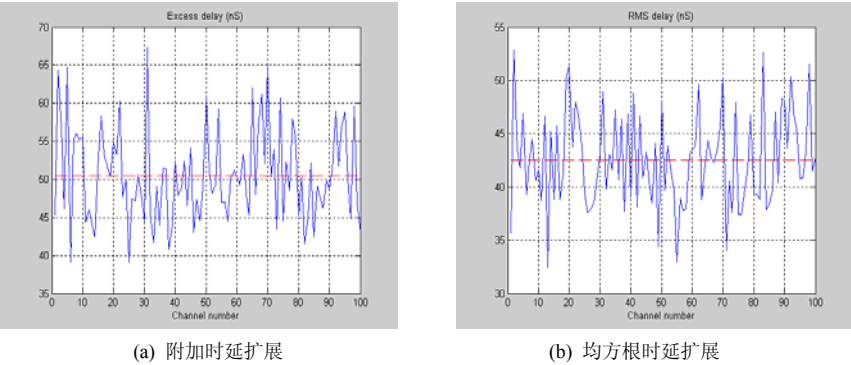


图 3 时延扩展特性

由图 3(a)和(b)看出，UWB 矿井巷道中存在时延为 40 纳秒以上的多径分量，这会在高速 UWB 通信系统中引入非常严重的码间干扰，因此信道模型的结构对系统设计有很大的影响。

2.2 MB-OFDM

MB-OFDM 是一种特殊的多载波调制（MCM）方式，它将一个较宽的频带划分为若干个子频带。

采用 MB-OFDM 调制解调方式，当子载波数较大时，各子载波幅度谱叠加的总信号的幅度谱有很好的矩阵特性，因此可以使 UWB 系统达到很高的频谱利用率。根据各个 UWB 设备之间或 UWB 设备与其它无线通信设备之间的干扰情况，以及不同矿井巷道的不同干扰源，可以降低一个或多个子频带的发射功率，或取消一个或多个子频带的发射，来有效降低干扰。还可以通过抑制相应的子载波，更加精确地控制合成信号的频谱形状。

2.3 仿真模型

建立的基于 MB-OFDM 的矿井 UWB 通信系统仿真模型，如图 4。OFDM 合成信号的产生方法与传统的 OFDM 系统类似，只是符号长度、子载波间隔、循环前缀长度等的具体参数与传统系统有较大差别。信号经天线发射出去，经过修正后的 UWB 信道，到达接收端。信号经过逆映射、解交织、解扰等处理，最后输出相应的数据。

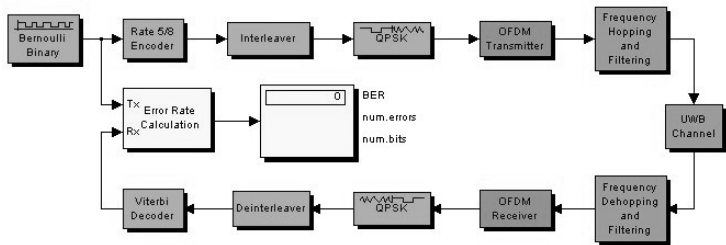


图 4 无线通信系统的 SIMULINK 仿真框图

3 性能仿真分析

3.1 性能指标

信号信噪比（SNR）的完全同步和符号间干扰是各种同步的结果，包括完全同步，不完全的定时同步，不完全的频率同步，以及不完全频率和时间同步。造成了信道间干扰（ICI）和符号间干扰（ISI）的产生。除了衰落和干扰，接收信号还受到 AWGN（加性高斯白噪声）的影响，其差异值是 N_0 。平均每 QPSK 符号的信噪比可以被界定为：

$$\overline{\gamma_s}(\varepsilon, \tau) = \frac{E_s \sigma_H^2}{\delta_C^2 + \delta_S^2 + N_0} \quad (4)$$

因为每比特功率 $E_b = E_s/2$ ，所以每比特的 SNR 平均为：

$$\overline{\gamma_b}(\varepsilon, \tau) = (1/2) \overline{\gamma_s}(\varepsilon, \tau) \quad (5)$$

另外，我们可以定义接收端的信噪比为 SNR：

$$\text{SNR}_\gamma = E_b / N_0 \quad (6)$$

这里 E_b 是接收到的每比特的平均功率， N_0 是加性高斯白噪声的单边带功率谱密度，即离散噪声。

平均误码率（BER）被定义为从整个发送端接收到的错误比特的比率，BER 是 OFDM 信道必然存在的方面。文献[8]中，使用了 $|H[k]|^2$ 的分布，我们也能获取 OFDM UWB 的平均 BER 的方案。信道模型中的 BER 可分为两类：Uncoded BER(UCBER)和 Coded BER(CBER)，他们的符号可以定义为：

$$UC_{\text{BER}} = E[Q(\sqrt{2|h|^2 \text{SNR}_\gamma})] = \frac{1}{2} \left(1 - \sqrt{\frac{\text{SNR}_\gamma}{1 + \text{SNR}_\gamma}}\right) \quad (7)$$

这里 Q 是标准正态分布的互补累计分布部分。

CBER，为经卷积译码器后被定义的部分。每个编码位有相同的概率被产生误码。卷积编码器是在一个理想的时频交织之后。

$$C_{\text{BER}} \leq \frac{\partial T(U, I)}{\partial I} \bigg|_{I=1, U=\sqrt{4UC_{\text{BER}}(1-C_{\text{BER}})}} \quad (8)$$

这里 T(U,I)是卷积编码器生成的函数，定义 d_{free} 为卷积码的最小自由距离，定义 N_b 为所有输入序列的汉明重量的总和，利用前面的推导过程， C_{BER} 可以表示为：

$$C_{\text{BER}} \approx N_b [4UC_{\text{BER}}(1-UC_{\text{BER}})]^{\frac{d_{\text{free}}}{2}} \quad (9)$$

因此，我们可以得出

$$C_{\text{BER}} \approx N_b [1 + \text{SNR}_\gamma]^{\frac{d_{\text{free}}}{2}} \quad (10)$$

3.2 仿真分析

在 200Mbps 速率模式，每帧 OFDM 符号数为 60 的情况下，通过增加有用数据子载波、提高编码速率或者减小循环前缀、保护间隔长度和训练序列长度来提高频带利用率。

3.2.1 传输能力

初始化信道的 SNR(信噪比)为一个较大的值(60dB)，以便减小干扰误差。发送端最终发送的和接收端接收到的 OFDM 信号的频谱图如图 5 所示。由图中可以看出，OFDM 信号具有很好的频率选择性和抗多径时延能力，连续三帧的 OFDM 信号分别占用三个不同的频带，每个频带的带宽大致都为 528MHz，与设计原理是相符的，且证明了系统能有效抵抗多径衰落。

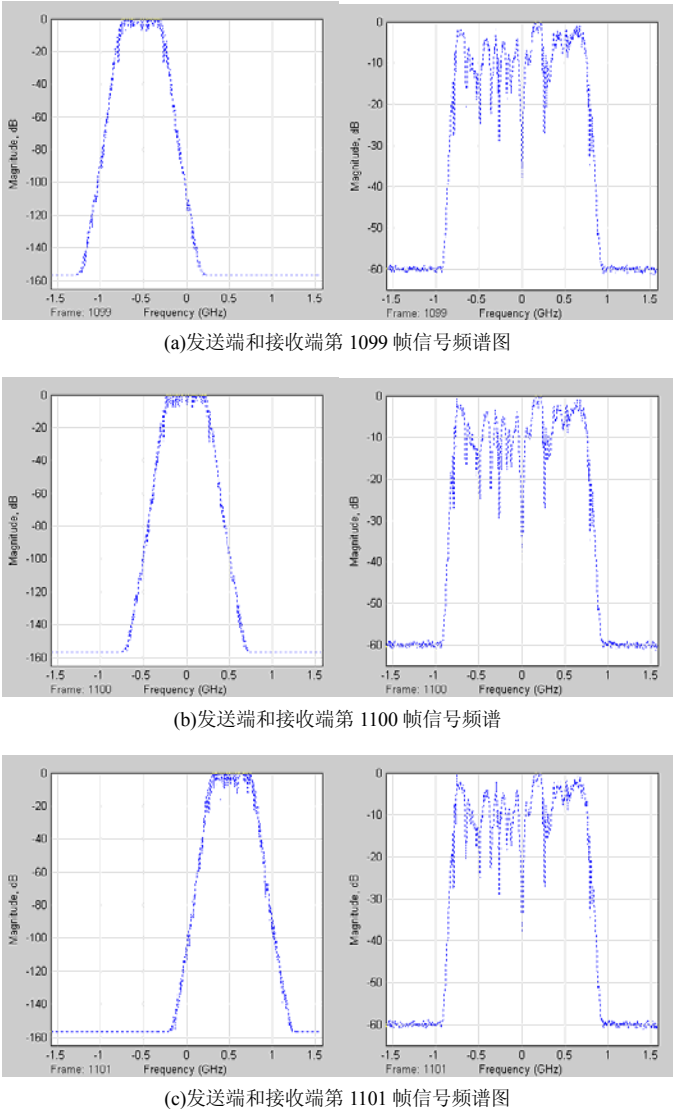


图 5 连续三帧发送和接收到的 OFDM 信号的频谱图

3.2.2 抗噪声能力

大量的信号从 OFDM 发射机中发射出来时，近似的按照多径频率选择衰落的特性来进行。OFDM 的动态范围为 30dB 左右，发射机发射的信号通过不同频带进行发射。图 6，在 UWB 信道环境下，虽然 SNR 取值很大，达到了 60dB，在没有经过信道估计的情况下，信号的幅度

和相位都发生了严重的偏移；然而，在经过信道估计和同步后，相位得到了完全的纠正，接收机信号分布图呈分布均匀的 X 状。

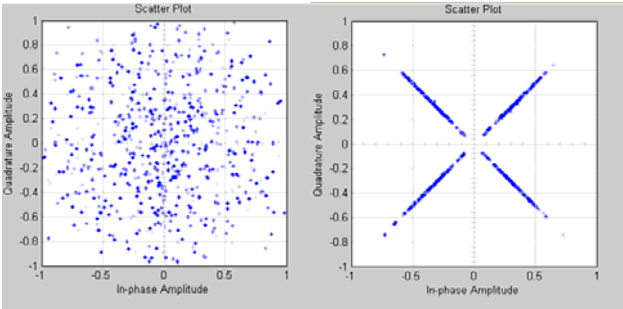


图 6 SNR=60dB 时，无信道估计和信道估计和同步后，接收信号的星座散列图

3.2.3 误码率

图 7 是误码率（BER）随信噪比（SNR）变化曲线，OFDM 在 QPSK 调制下，对各个参数进行仿真。OFDM-UWB 无线系统的子带带宽为 528MHz，信噪比范围 8~17db。

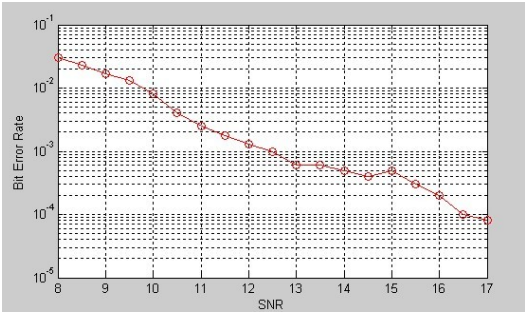


图 7 系统误码率随信噪比变化的曲线

信噪比提高时，误码率曲线下降。在 14db 和 15db 之间有一个转折点。

3.2.4 覆盖范围

系统的数据传输速率分别为：110Mbps、480Mbps 时，信道可通率不小于 90%，系统的传输距离与误包率（PER）的关系，如图 8：

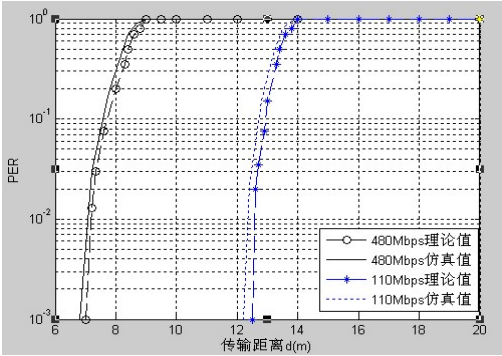


图 8 110Mbps、480Mbps 时，PER 随传输距离变换曲线

可以看出，仿真结果与理论值基本吻合，仿真与理论值差异的根本原因是信道系数 $h(t)$ 的分布不同。

4 结束语

本文把 MB-OFDM 抗多径时延的特性与修正的信道损耗模型相结合，建立基于 MB-OFDM 的矿井 UWB 无线通信系统模型。通过仿真表明，系统能有效抵抗多径衰落，具有抗噪声能力强，误码率低，覆盖范围远等优点。

参 考 文 献

- [1] 孙继平, 石庆冬. 矿井隧道电磁传播的研究[J]. 煤炭科学技术, 2001 (1).
- [2] 杨维, 冯锡生, 孙继平等. 新一代全矿井无线信息系统理论与关键技术[J].煤炭学报, 2004 (4) .
- [3] Moe Z W, George C, Nelson R S. Rerformance of rake reception in dense multipath channels: Implications of spreading band-width and selection diversity order[J]. IEEE Journal on Selected Areas in Communications, 2000 (8).
- [4] 彭丽, 孙彦景, 钱建生等. UWB 矿井巷道无线信道模型[J]. 计算机工程与设计, 2009(3).
- [5] 孙继平, 苏又平, 梁良. 900MHz DDS/PLL 在矿井无线通信系统中的应用[J]. 煤炭科学技术, 2001 (10).
- [6] 张跃平, 张文梅, 盛剑桓等. 900MHz 无线电波在地下长壁煤矿的传播[J]. 煤炭学报, 2002 (1) .
- [7] 李长青. UWB-OFDM 系统中若干关键技术的研究 [D]. 北京: 北京邮电大学, 2006.
- [8] Qiyue Zou, Alireza Tarighat, Ali H. Sayed. Performance Analysis of Multi-band OFDM UWB Communications With Application to Range Improvement[J].IEEE Transactions on Vehicular Technology, 2007 (11)

作者简介:

钱建生 (1964-), 男, 汉族, 浙江乌镇, 博士, 中国矿业大学, 教授, 矿井通信与监控;
王 鑫 (1984-), 男, 汉族, 山东烟台, 硕士生, 中国矿业大学, 矿井无线通信信道模型;
孙彦景 (1977-), 男, 汉族, 山东滕州, 博士, 中国矿业大学, 副教授, 无线感知网络;

序列相关性对系统性能影响的理论分析和评估

吴 娜 吴晓非 邹永忠

(北京邮电大学 信息与通信工程学院, 北京 100876)

摘 要: 本文基于序列的相关特性, 推导了 CDMA 系统的理论性能, 并在 AWGN 信道和单径及多径 Rayleigh 信道模型下, 对 GOLD 码和 LS 码进行性能评估和比较。

关键词: 通信技术; LS 码/ GOLD 码; 性能评估; CDMA

The Theoretic Analysis and Evaluation of Sequences Correlation Affecting the Systems' Performance

Wu Na Wu Xiao fei Zou Yong zhong

(Beijing University of Posts and Telecommunications, Beijing 100876, China)

Abstract: Based on the Sequences correlation, in this paper ,the theoretic performance of CDMA system is deduced, then the performance evaluation and comparison of GOLD code and LS code are done, under AWGN ,singlepath and multipath Rayleigh channel model respectively.

Keywords: Communication technology ;LS code/Gold code ; the performance evaluation ; CDMA

1 引言

码分多址 (CDMA Code Division Multiple Access) 是一种以扩频通信为基础的载波调制和多址接入技术。在 CDMA 系统,所有用户使用同一频带,系统在发送端使用特定扩频码对信号进行扩频,在接收端利用扩频码的相关性来区分不同用户。因此码的设计好坏直接影响系统性能。

CDMA 系统的干扰可分为四大部分: 噪声、符号间干扰 (ISI Intersymbol interference)、多址干扰 (MAI Multiple access interference) 和相邻信道间干扰 (ACI Adjacent channel interference)。对 CDMA 系统而言, ISI, MAI, ACI 均可以靠选择性能良好的扩频码使之降低甚至消除。

为了减小 ISI, 每个用户所选用的地址码的自相关函数应该是一个理想的冲激函数, 即除原点外, 应处处为零。为了减小 MAI, 各个用户所选用的地址码间的互相关函数应处处为零。从正交性的观点来讲, 各个地址码与其自身除相对零时延处外, 对任何非零相对时延都应该相互正交, 而地址码间对任何相对时延 (含零时延) 都应相互正交。

遗憾的是, 已经证明具有这种理想相关特性的扩频序列集是不存在的。对给定的序列周期与数目, 一个序列集最大自相关函数副峰值和最大互相关函数值不可能同时为零, 它们受到理论界 (如 Welch 界, Sidelnikov 界, Sarwate 界^[1-4]等) 的约束, 一个变小时, 另一个必然

增大。所以实际系统中只能选用具有较好相关特性的扩频码，如 Gold 码等^{[7][8]}，但多址干扰始终是存在的。近年来，扩频码设计取得了重要突破，李道本教授提出了能消除 MAI 及 ISI 的具有零相关窗的 LS (Loosely Synchronized) 序列^[6]；范平志教授提出了广义正交 (GO Generalized Orthogonality) 序列^[5]的概念及 LCZ (Low correlation zone) 序列。

本文内容组织如下：第二部分介绍系统模型；第三部分理论分析扩频码的相关性对系统性能的影响；第四部分为性能评估的仿真结果；第五部分小结。

2 系统模型

考察单小区 DS-CDMA 系统前向链路，假设有 K 个用户和 L 径，每个用户分配一个持续时间为 T 的特征波形 $a_k(t)$ ，其中 T 为符号间隔， $T = NT_c$ 。特征波形 $a_k(t)$ 为

$$a_k(t) = \sum_{n=0}^{N-1} a_{(n)}^{(k)} p(t - nT_c) \quad , \quad k=1,2,\dots, K \quad (1)$$

其中 T_c 是码片间隔， N 是序列长度， $p(t)$ 是宽度为 T_c 的成形脉冲，如果没有特殊说明， $p(t)$ 为矩形方波。不失一般性，假设所有的 K 个特征波形 $a_k(t)$ 具有单位能量。

第 k 个用户的信道 $h_k(t)$ 的等效的低通的离散表示为：

$$h_k(t) = \sum_{l=0}^{L_k-1} \alpha_{k,l} \delta(t - \tau_{k,l}) \exp(j\phi_{k,l}) \quad (2)$$

其中 $\alpha_{k,l}$ 是表示第 k 个用户第 l 径的幅度衰减系数(增益)， $\phi_{k,l}$ 表示相位偏移， $\tau_{k,l}$ 是第 l 径时延，不妨假设有 $0 \leq \tau_{k,0} < \tau_{k,1} < \dots < \tau_{k,L_{k-1}} < T$ 。

$s^{(k)}(t) = \sum_{j=0}^{\infty} s_{(j)}^{(k)} \text{Rect}(t - jT)$ 是第 k 个用户的数据符号， $s_{(j)}^{(k)}$ 表示第 k 个用户的在时间段 $[jT, (j+1)T]$ 传送的第 j 个数据符号，考虑到功率分配和调整，进一步假设第 k 个用户的所有传输信号都具有的发射功率调整因子 P_k ，对应于幅度调整因子为 $\sqrt{P_k}$ 。

接收端接收到的等效的基带信号可以表示为

$$r(t) = \sum_{k=1}^K \sum_{l=0}^{L-1} \sqrt{P_k} \alpha_{k,l} \exp(j\phi_{k,l}) s^{(k)}(t - \tau_{k,l}) a_k(t - \tau_{k,l}) + n(t) \quad (3)$$

$n(t)$ 是双边功率谱密度为 $N_0/2$ 的加性白高斯噪声 (AWGN)。

3 理论性能分析

3.1 相关特性分析

假设参数 $\alpha_{k,l}$ ， $\phi_{k,l}$ ， $\tau_{k,l}$ 的估计是理想的，将解扩后的 $r_{j,n}^{(i)}$ 能量 (功率) 最大的 B 径 ($B \leq L_i$) 采用最大比值合并 (MRC, Maximum Ratio Combining) 接收，不妨假设前 B 径即为能量 (功率) 最大的 B 径，那么第 i 个用户的第 j 个符号的估计 $r_j^{(i)}$ ，

$$r_j^{(i)} = \sqrt{P_i} s_{(j)}^{(i)} \sum_{l=0}^{B-1} \alpha_{i,l}^2 + I_{ISI}^{(i,j)} + I_{MAI}^{(i,j)} + \eta^{(i,j)} \quad (4)$$

第一项表示第 i 个用户的第 j 个符号的第 n 径解扩后的期望信号；第二项 $I_{ISI}^{(i,j)}$ 表示第 i 个用户的第 j 个符号的最大比值合并后受到的多径自干扰大小；第三项 $I_{MAI}^{(i,j)}$ 表示表示第 i 个用户的第 j 个符号的最大比值合并后受到的多用户的干扰；第四项表示第 i 个用户的第 j 个符号的最大比值合并后受到的噪声大小。

$$I_{ISI}^{(i,j)} = \sum_{n=0}^{B-1} \left\{ \sum_{l=0, l \neq n}^{L < n} \sqrt{P_i} \alpha_{i,n} \alpha_{i,l} \exp(j\phi_{i,l} - j\phi_{i,n}) [s_{(j)}^{(i)} \rho_{i,i}^* (\tau_{i,n} - \tau_{i,l}) + s_{(j+1)}^{(i)} \hat{\rho}_{i,i}^* (\tau_{i,n} - \tau_{i,l})] + \right. \\ \left. \sum_{l=n+1}^{L_i} \sqrt{P_i} \alpha_{i,n} \alpha_{i,l} \exp(j\phi_{i,l} - j\phi_{i,n}) [s_{(j-1)}^{(i)} \hat{\rho}_{i,i}^* (\tau_{i,n} - \tau_{i,l}) + s_{(j)}^{(i)} \rho_{i,i}^* (\tau_{i,n} - \tau_{i,l})] \right\} \quad (5)$$

$$I_{MAI}^{(i,j)} = \sum_{k=1, k \neq i}^K \left\{ \sum_{\tau_{k,l} \leq \tau_{i,n}} \sqrt{P_k} \alpha_{i,n} \alpha_{k,l} \exp(j\phi_{k,l} - j\phi_{i,n}) [s_{(j)}^{(k)} \rho_{i,k}^* (\tau_{i,n} - \tau_{k,l}) + s_{(j+1)}^{(k)} \hat{\rho}_{i,k}^* (\tau_{i,n} - \tau_{k,l})] + \right. \\ \left. \sum_{\tau_{k,l} > \tau_{i,n}} \sqrt{P_k} \alpha_{i,n} \alpha_{k,l} \exp(j\phi_{k,l} - j\phi_{i,n}) [s_{(j-1)}^{(k)} \hat{\rho}_{i,k}^* (\tau_{i,n} - \tau_{k,l}) + s_{(j)}^{(k)} \rho_{i,k}^* (\tau_{i,n} - \tau_{k,l})] \right\} \quad (6)$$

$$\eta^{(i,j)} = \sum_{n=0}^{B-1} \alpha_{i,n} \exp(-j\phi_{i,n}) \eta^{(i,j,n)} \quad (7)$$

这里用到了非周期相关函数的定义^[9]，假设 $mT_c \leq \tau < (m+1)T_c$ ， m 为正整数。

$$\hat{\rho}_{k,i}(\tau) = \frac{1}{T} \{ [(m+1)T_c - \tau] C_{k,i}(m-N) + (\tau - mT_c) C_{k,i}(m+1-N) \}, \quad 0 \leq \tau < T \quad (8)$$

$$\rho_{k,i}(\tau) = \frac{1}{T} \{ [(m+1)T_c - \tau] C_{k,i}(m) + (\tau - mT_c) C_{k,i}(m+1) \}, \quad 0 \leq \tau < T \quad (9)$$

$$\hat{\rho}'_{k,i}(\tau) = \frac{1}{T} \{ [(m+1)T_c - \tau] C_{k,i}(m+N) + (\tau - mT_c) C_{k,i}(m+1+N) \}, \quad 0 \leq \tau < T \quad (10)$$

$$\rho'_{k,i}(\tau) = \frac{1}{T} \{ (\tau - mT_c) C_{k,i}(m+1) + [(m+1)T_c - \tau] C_{k,i}(m) \}, \quad 0 \leq \tau < T \quad (11)$$

其中的 $C_{k,i}(m)$ 为非周期互相关函数，定义如下：

$$C_{k,i}(m) = \begin{cases} \sum_{n=0}^{N-m-1} a_{(n)}^{(k)} a_{(n+m)}^{(i)*}, & 0 \leq m \leq N-1 \\ \sum_{n=0}^{N+\tau-1} a_{(n-\tau)}^{(k)} a_{(n)}^{(i)*}, & 1-N \leq m \leq 0 \\ 0, & |m| \geq N \end{cases} \quad (12)$$

根据 $\hat{\rho}_{k,i}(\tau)$ 、 $\rho_{k,i}(\tau)$ 、 $\hat{\rho}'_{k,i}(\tau)$ 、 $\rho'_{k,i}(\tau)$ 的均值和方差，假设 $(m+1)T_c - \tau$ 服从 $[0, T_c]$ 的均匀分布， $E[s_{(j)}^{(k)}] = 0$ ， $E[(s_{(j)}^{(k)})^2] = 1.0$ ， $\phi_{k,l}$ 在 $[0, 2\pi)$ 上均匀分布，进一步假设对于指定的 $s_{(j)}^{(i)}$ ， P_k ， $\tau_{k,l}$ ， $\alpha_{k,l}$ ， $\phi_{k,l}$ 是常数，随机变量 $s_{(j)}^{(k)}$ ， $\tau_{k,l}$ ， $\alpha_{k,l}$ ， $\phi_{k,l}$ 相互独立，可以推导出 $I_{ISI}^{(i,j)}$ 、 $I_{MAI}^{(i,j)}$ 的一阶和二阶统计特性。由于扩频码的非周期相关函数确定了干扰的大小，为了考察干扰值和互干扰值，引进如下定义：

$$Y_{k,i}(m) = C_{k,i}^2(m - N) + C_{k,i}^2(m), \quad 0 \leq m < N, m \text{ 为整数} \quad (13)$$

$$Y'_{k,i}(m) = C_{k,i}^2(m + N) + C_{k,i}^2(m), \quad -N < m < 0, m \text{ 为整数} \quad (14)$$

令 $-T < m_{n,l}^{(i,k)} T_c \leq \tau_{i,n} - \tau_{i,l} < (m_{n,l}^{(i,k)} + 1)T_c \leq T$, $m_{n,l}^{(i,k)}$ 为整数, 可得到如下公式:

$$Var[I_{ISI}^{(i,j)}] = \frac{T_c^2 P_i}{T^2} \sum_{n=0}^{B-1} \alpha_{i,n}^2 \left\{ \sum_{l=0, l \neq n}^{l < n} E[\alpha_{i,l}^2] Y_{i,i}(m_{n,l}^{(i,k)}) + \sum_{l=n+1}^{L_i} E[\alpha_{i,l}^2] Y'_{i,i}(m_{n,l}^{(i,k)}) \right\} \quad (15)$$

$$Var[I_{MAI}^{(i,j)}] = \frac{T_c^2}{T^2} \sum_{n=0}^{B-1} \alpha_{i,n}^2 \left\{ \sum_{k=1, k \neq i}^K \left\{ \sum_{\tau_{k,l} \leq \tau_{i,n}} P_k E[\alpha_{k,l}^2] Y_{k,i}(m_{n,l}^{(i,k)}) + \sum_{\tau_{k,l} > \tau_{i,n}} P_k E[\alpha_{k,l}^2] Y'_{k,i}(m_{n,l}^{(i,k)}) \right\} \right\} \quad (16)$$

3.2 SINR分析

假设 P_k 是常数, $\alpha_{k,l}$, $\varphi_{k,l}$, $\tau_{k,l}$ 的估计是理想的, 所有的用户发射功率调整因子相同, 不妨将其归一化为 1。将上述的 $r_{j,n}^{(i)}$ 将能量 (功率) 最大的 B 径 ($B \leq L_i$) 采用最大比值合并 (MRC, Maximum Ratio Combining) 接收, 不妨假设前 B 径即为能量 (功率) 最大的 B 径, 那么第 i 个用户的第 j 个符号的估计 $r_j^{(i)}$ 的信号与干扰和噪声比 (SINR Signal-Interference-Noise-Ratio) 可以推导如下:

$$SINR_{av} = \frac{\left(\sum_{l=0}^{B-1} \alpha_{i,l}^2 \right)^2}{\left(\frac{N0}{2NE_s} \sum_{l=0}^{B-1} \alpha_{i,l}^2 + \frac{1}{3N^2} \sum_{n=0}^{B-1} \alpha_{i,n}^2 \left\{ \sum_{l=0, l \neq n}^{L-1} E[\alpha_{i,l}^2] Y_{i,i}(m_{n,l}) + \sum_{l=n+1}^{L_i} E[\alpha_{i,l}^2] Y'_{i,i}(m_{n,l}) \right\} + \frac{1}{3N^2} \sum_{n=0}^{B-1} \alpha_{i,n}^2 \left\{ \sum_{k=1, k \neq i}^K \left\{ \sum_{l=0}^{n-1} E[\alpha_{k,l}^2] Y_{k,i}(m_{n,l}) + \sum_{l=n+1}^{L-1} E[\alpha_{k,l}^2] Y'_{k,i}(m_{n,l}) \right\} \right\} \right)} \quad (17)$$

4 性能评估及仿真结果

CDMA 系统中可以对于干扰采用高斯近似的方法 [9][10], 其理论依据是中心极限定理, 文章[11]指出当 SNR 较高时, 干扰采用高斯近似能保证足够高的精度, 文中如果没有特别说明, 对于所有的干扰都采用高斯近似来推导系统性能。

注意, 文中的所有仿真均将扩频处理增益作为系统的增益, 也就是说, 系统所需要的实际射频前端的信噪比比仿真的信噪比可以低 18dB(64 倍扩频的处理增益是 18dB)。

仿真环境如下: $K=2$, $N=64$, $L=6$, 假设 P_k 是常数, 所有用户共享相同的 6 个多径信道, 相应的参数 $\tau_{k,l}$, $a_{k,l}$, $\varphi_{k,l}$ 都相同, 而且其估计是理想的, 令 $\tau_{k,l}=0$, $\tau_{k,l}=mT_c$, m 为正整数, $T_c=1953\text{ns}$ 。

4.1 AWGN信道下BER的性能

LS 码的方案比 GOLD 码方案有约 0.5dB 的性能增益, 仿真结果如图 1 所示。增益产生的原因在于 GOLD 码由基带成型滤波器带来的干扰较大, 如果不限带的话, 二者的性能是基本一致的。

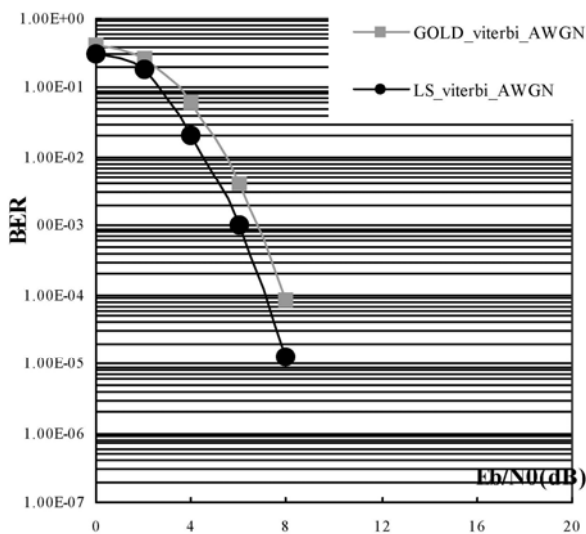


图1 AWGN 信道下 GOLD 码和 LS 码性能曲线比较

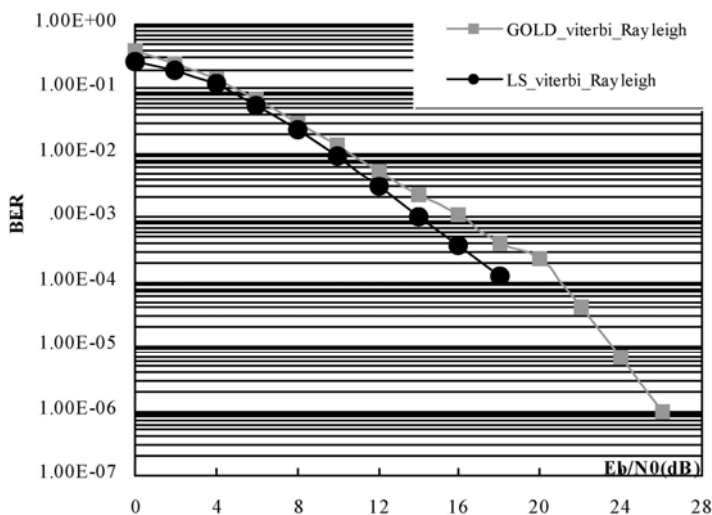


图2 单径 Rayleigh 信道下 Gold 码和 LS 码性能曲线比较

4.2 单径Rayleigh信道下BER的性能

采用理想信道估计，载频 450MHz，步行速度 5km/h，1/2 卷积码,viterbi 译码，采用 64 位 LS 码和 GOLD 码扩频的性能比较，仿真结果如图 2 所示：

由仿真结果可知，GOLD 码方案在单径 Rayleigh 信道下可以正常工作，但是性能比 LS 码方案差，随着信噪比的增大，误码率差距增大，在 20dB 时采用 LS 码的方案已经没有误码，而 GOLD 码方案在 26dB 时 BER 为 10^{-6} ，因此对于语音业务来说，LS 码的方案比 GOLD 码方案仅有 0.5~1.0dB 的增益(误码率为 $10^{-2} \sim 10^{-3}$ 即可)，但是对于数据/图象业务来说（BER 一般要求低于 10^{-6} ），LS 码的方案比 GOLD 码方案有 6dB 以上的增益。

4.3 多径Rayleigh信道下BER的性能

在多径 Rayleigh 信道下，载频 450MHz，步行速度 5km/h，采用理想信道估计，QPSK 调制，1/2 卷积码，viterbi 译码，采用 64 位 LS 码扩频和 GOLD 码扩频的性能比较，仿真结果如图 3 所示：

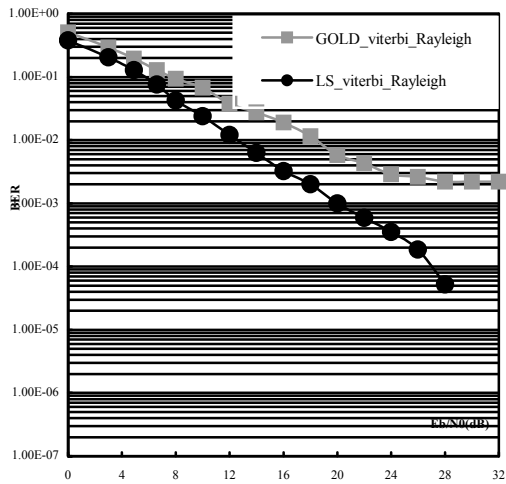


图 3 多径 Rayleigh 信道下 GOLD 码和 LS 码性能曲线比较

由仿真结果可知，GOLD 码方案在多径 Rayleigh 信道下难以正常工作，在误码率高于 10^{-3} 是已经出现误码平台，而 LS 码方案仍能正常工作，但是性能比 LS 码方案差，随着信噪比的增大，误码率差距增大，在 30dB 时采用 LS 码的方案已经没有误码，因此对于语音业务来说，LS 码的方案比 GOLD 码方案仅有 6.0dB 的增益(误码率为 $10^{-2} \sim 10^{-3}$ 即可)，但是对于数据/图象业务来说（BER 一般要求低于 10^{-6} ），GOLD 码方案已经不能正常工作，LS 码的方案仍然可以工作得很好。

5 小结

本文通过对扩频码的相关特性进行统计分析，在此基础上得出 ISI 和 MAI 的统计分析，进而得到系统在 AWGN 信道下、Rayleigh 单径及多径信道下的误码率曲线。由理论推导和仿真结果可以得出，扩频序列的优化是 CDMA 系统性能提高的核心问题。

参 考 文 献

[1] L. R. Welch. Lower bounds on the maximum crosscorrelation of signals, IEEE Trans. Inform. Theory.1974,397~399

[2] D. V. Sarwate. Bounds on crosscorrelation and autocorrelation of sequences, IEEE Trans. Inform. Theory.1979,25:720~724

[3] V. M. Sidelnikov. On mutual correlation of sequences, Soviet Math Doklady.1971,12:197~201

- [4] J. L. Massey. On Welch's Bound for the crosscorrelation of a sequence set, Proceedings of IEEE ISIT'90.1990,385
- [5] P. Z. Fan, N. Suehiro, N. Kuroyanagi and X.M. Deng. A class of binary sequences with zero correlation zone, IEE Electron. Lett..1999,35(10):777~779
- [6] D.B., Li, High spectrum efficient multiple access code, Chinese Journal of Electronics, vol. 8, no. 3, pp. 221-226, 1999.
- [7] P. Z. Fan and M. Darnell. 《Sequence Design for Communications Applications》. Research Studies Press, John Wiley & Sons Ltd, London, 1996. ISBN 0-471-96557-x, 516 pages.
- [8] P. Z. Fan and M. Darnell. On the construction and comparison of periodic digital sequence sets. IEE Proceedings Communications. 1997,144(6):111~117
- [9] Pingzhi Fan and Li Hao, Generalized orthogonal sequences and their applications in synchronous CDMA systems, IEICE TRANS. FUNDMENTALS, VOL.E-83-A, NO.11 November 2000.
- [10] M.B. Pursley, "Performance evaluation for phase-coded spread spectrum multiple-access communications—Part I: System analysis," IEEE Trans. Comm., vol. COM-25, pp.795-799, Aug. 1977.
- [11] C. Trabelsi and A. Yongacoglu, "Bit-error-rate performance for asynchronous DS-CDMA over fading channels," IEE Proc. Comm. Vol.142, no.5, pp. 307-314, Oct. 1995.

基于中继节点的无线传感器网络分簇路由算法

李 雷¹ 付东阳² 李艳辉²

(1. 南京邮电大学 理学院, 江苏 南京 210003;

2. 南京邮电大学 自动化学院, 江苏 南京 21003)

摘 要: 无线传感器网络(Wireless Sensor Networks 简称 WSN)节点数量庞大、单个节点资源极其有限, 其路由协议设计的首要目标是有效节约能源、延长网络生命周期。LEACH(Low-Energy Adaptive Clustering Hierarchy)是层次路由协议中的经典协议, 大多数层次路由协议以这种协议为基础。在分析了 LEACH 的工作原理和优缺点后, 提出了一种新的基于中继节点的分簇路由协议 CPRN(Clustering Protocol based on Relay Nodes)。协议首先使用 LEACH 中的方法对节点进行分簇并选出簇头, 然后再通过网络的能量感知性选出中继节点。簇形成后, 簇头和中继节点相互合作将数据传到基站。仿真结果显示, CPRN 协议能有效的均衡能量负载, 延长网络生命周期。

关键词: 通信; 能量负载均衡; 分簇; 中继节点; 无线传感器网络

A Clustering Routing Approach with Relay Nodes in Wireless Sensor Networks

Li Lei¹ Fu Dong yang² Li Yan hui²

(1.2. Pattern Recognition and Intelligence System, Nanjing University of Posts and Telecommunications, Nanjing 21003, China)

Abstract: The number of sensor nodes in WSN is numerous and single node is extraordinarily limited in resource, so the important aim of designing routing protocol of WSN is to reduce the overall energy-dissipated in the networks and to maximize the networks lifetime. LEACH is typical hierarchical protocol which is the foundation of most hierarchical protocols. Most hierarchical routing protocol based on this protocol. After analyzing the protocols on theory, advantages and disadvantages, this paper proposes a new clustering protocol on relay nodes (CPRN). Firstly, this protocol builds up clusters and selects cluster head with the same method as LEACH, then through the network of energy-aware it selects relay nodes. After cluster formation, Cluster Heads and relay nodes cooperatively transmit information to the BS. Simulation results show that comparing with LEACH, CPRN can balance energy load and prolongs the lifetime of WSN.

Keywords: Communication; Energy Load Balancing; Clustering; Relay Nodes; Wireless Sensor Network

0 引言

无线传感器网络^[1]是由部署在监测区域内大量的廉价微型传感器节点组成,通过无线通信方式形成的一个多跳的自组织网络系统,其目的是感知、采集和处理网络覆盖区域中感知对象的信息,并发送给观察者。无线传感器网络具有十分广阔的应用前景,在军事国防、工农业、城市管理、生物医疗、环境监测、抢险救灾、危险区域、远程控制等许多领域都有重要的科研价值和巨大实用价值^[2]。

网络数据传输离不开路由协议。分簇算法是目前最有效的 WSN 路由协议。通过约定的协议,在整个网络覆盖范围内选举出一定数量的节点作为簇头,从而将网络划分为簇,簇头管理簇内的普通节点,簇头通过信息的交互完成全网拓扑的构建。目前已出现大量的分簇路由协议,从不同的角度来延长传感器网络的寿命。而 LEACH 协议[3]是最早提出的分簇路由协议。它的基本思想是:通过等概率的随机循环选择簇头,将整个网络的能量负载平均分配到每个传感器节点,从而达到降低网络能量耗费,延长网络生命时间的目的。

LEACH 协议只考虑了单跳模型,不适合大型 WSN 网络。文献[4]提出 M-LEACH(multi-hop variant of LEACH,多跳 LEACH 算法)。簇内的节点不再是以单跳的方式传输数据到簇头,而是通过簇内其他节点转发。文献[5]提出 EEUC (energy-efficient unequal clustering 能量高效的非均等分簇)算法。它利用非均匀的竞争半径,使得靠近汇聚点的簇的成员数目相对较小,从而簇头能够节约能量以供数据转发使用,达到均衡簇头能量消耗的目的。此外,在簇头选择单个节点时,不仅考虑候选节点相对汇聚点的位置,还应考虑候选节点的剩余能量。

本文的 CPRN 协议通过使簇头和中继节点协作通信来降低无线通信模块的能耗,同时将网络划分为一些互不重叠的区域。簇头位于每个区域的中央。它通过能量感知来选出中继节点实现负载均衡,可有效的降低自身的能量消耗。簇头是网络中最重要的节点,一旦簇头能耗殆尽,则会导致网络瘫痪,所以本协议建立了一个很好的检测系统来监视簇头节点的状态。而这个检测工作是靠簇内所有的成员节点共同完成的。仿真结果显示 CPRN 是一个低开销,能量有效的分簇路由算法。

1 算法描述

1.1 系统能量模型

本文采用如下的传感器模型: N 个传感器节点随机均匀分布在一个正方形区域内,我们可以做一下假设:

- (1) 在区域外很远的一个固定位置有唯一的基站;
- (2) 网络中所有节点都是同构且能量有限的;
- (3) 节点可根据接受信号的强度来判断基站相对自己的远近;
- (4) 无线发射功率可控,即节点可以根据距离来调整发射功率的大小。

我们采用一个简化的硬件能量消耗模型^[6]。传感器节点发送和接受 1 bit 信息,距离为 d 时,所消耗的能量分别为:

$$E_{Tx}(l, d) = E_{Tx\text{-elec}}(l) + E_{Tx\text{-amp}}(l, d)$$

$$= \begin{cases} lE_{\text{elec}} + l_{\varepsilon_{fs}} d^2 & d < d_0 \\ lE_{\text{elec}} + l_{\varepsilon_{mp}} d^4 & d \geq d_0 \end{cases} \quad (1)$$

$$E_T(l) = E_{R\text{-elec}}(l) = lE_{\text{elec}} \quad (2)$$

在式(1)中, 如果接受、发送器之间的距离小于临界值 d_0 , 则使用自由空间模型; 如果大于临界值 d_0 , 则使用多路径模型。在(2)中, E_{elec} 是发射电路和接受电路所消耗的能量, 它取决于电路的数字编码、调制和滤波等因素, 在这个模型中发射与接受电路两者相同。

1.2 CPRN算法详述

CPRN 协议将网络生命周期划分为很多时间固定的轮数。每一轮分为簇的建立阶段和稳定的通信阶段。在簇的建立阶段, 节点通过相互竞争成为簇头。这个阶段又分为三个子阶段。在稳定阶段, 为了避免传输数据时发生碰撞, 簇内的普通节点采用可变长度的 *TDMA* 时隙在各自的时间槽内发送数据。

1.2.1 簇头选举及簇结构的形成

我们使用 *LEACH* 来选举簇头。节点是否成为簇头取决于网络中事先假定成为簇头节点的百分比和到目前为止节点成为簇头的次数。通过设定一个阈值 $T(n)$, 并让节点在 0 — 1 之间选择一个随机数, 如果该随机数比 $T(n)$ 小, 则该节点在当前轮成为簇头。当选过的节点的 $T(n)$ 设置为 0, 这样该节点就不会再次当选为簇头。随着簇头的节点数增加, $T(n)$ 随之增大, 当剩下一个节点未当选时 $T(n)$ 为 1, 表示这个节点一定当选。 $T(n)$ 可表示为

$$T(s) = \begin{cases} \frac{p_{opt}}{1 - p_{opt} \cdot \left(r \bmod \frac{1}{p_{opt}} \right)} & \text{if } s \in G \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

其中, P 是簇头在所有节点中占得百分比, r 是当前轮数, G 是在剩余的 $1/p$ 轮种成为簇头节点的节点集合。

LEACH 簇头选举算法没有考虑节点的具体位置, 不能保证簇头在网络中的均匀分布, 由此而形成簇与簇之间相互重叠以及不对称分布。在 *CPRN* 协议中增加了一个校正阶段来弥补 *LEACH* 协议的不足。它将簇的形成动态的划分为三个阶段, 同时把竞争簇头的节点之间的距离也作为参考因素, 从而形成可靠的, 均匀分布的簇。

在形成簇的第一个阶段, 竞争节点使用间断的 *CSMA*(Carrier-Sense Multiple Access) *Mac* 协议来广播一个 *ADV* 报文。这个报文包含每个节点的 *ID*, 而且非常短小。竞争节点必须有足够的能量以支撑这个发射信号, 从而能使信号能够到达网络中的每个节点。周围节点根据接收到的簇头广播信号的强弱计算出和竞争节点之间的距离, 然后节点将这个距离和距离门限进行比较以决定加入哪个簇以及下一步做什么。根据网络的覆盖面积可得出这个距离门限值。网络的最优簇头数可有[7]得出。如果另一个竞争节点与发送者之间的距离小于距离门限值, 则该节点自动放弃竞争。也就是说如果两三个竞争节点在距离门限内, 相隔非常近, 则

他们都放弃竞争权利，以使发送邀请报文的节点成为簇头。CSMA MAC 协议中随机延迟时间与节点的剩余能量成反比，这可以保证剩余能量多的节点当选为簇头的机会更多。

在形成簇的第二个阶段，如果一个普通节点在多轮竞争之后，既没有当选为簇头也没有加入簇，它将会要求成为簇头，并发送邀请报文。CPRN 协议的目标就是建立一个簇分布均匀的网络拓扑，即使节点位于网络的边缘地带，也能形成合适的簇头和簇。最终，CPRN 遵循第一阶段的 CSMA MAC 协议，将网络划分为一定数量的互不重叠的簇。

在形成簇的第三个阶段，非簇头节点根据接收到的信号强度决定加入哪个簇。这些节点向簇头发布一个 JOIN 报文，这个报文也包含了节点当前的剩余能量。簇头作为控制中心协调簇内的数据传输。当它接受到所有的加入消息后，就产生一个 TDMA 定时消息，并通知簇内所有节点。簇头可以决定本簇内所有节点的 CDMA 编码。在通信阶段，簇内节点就使用这个编码向簇头传输数据。这些考虑确保了在数据传输时不存在任何冲突，也让每个非簇头节点的无线电组件除了在预先指定的时段打开其他时间一直处于关闭状态，大大节省节点能量。

1.2.2 簇的稳定路由通信

簇形成之后就进入稳定通信阶段。这个阶段也分为两个阶段：簇内数据传输和簇头与中继节点的协作传输。在前一阶段，有的成员节点处于休眠状态，一直到轮到它传输数据的时隙到来。簇内节点向簇头传送的数据包里面包含自己的 ID，采集的信息以及自身的剩余能量。此阶段的传输时间由簇内节点数目决定。在协作传输阶段，簇头将融合的数据广播给簇内的成员节点，其中剩余能量最多的节点再将此数据传向基站，这个节点我们成为中继节点。这样可以大大节省簇头的能量。我们注意到簇头向簇内成员广播信息可能会增加一些开销，然而这些广播仅在短距离的簇内广播，所以这些开销是可以忽略的。

1.2.3 系统的能量负载分布及鲁棒性

网络能量负载的均匀分布是系统的理想状态。在传统网络里面，簇头要比普通节点花费更多的能量。这就要求网络节点轮流当选为簇头。而簇头的再选择必然花费更多的开销。在 CPRN 协议中，我们引进的簇头和中继节点的协作传输，均衡了簇头的能量消耗以及簇内成员的能量负载。这种启发式的能量机制可以大大延长网络的生命周期。

簇头需要完成数据融合、与汇聚节点通信等工作，所以能量消耗大。一旦簇头能量耗尽，则整个簇都会瘫痪。为了解决这个问题，我们建立了一个节点故障检测系统，监视簇头节点的变化，以增加无线传感器网络的鲁棒性。在协作通信阶段，簇内所有节点都处于苏醒状态，准备接受来自簇头融合后的数据包。如果簇内成员在连续两个时隙后都没有收到簇头的数据包，那么就认为簇头出现故障。此时这些节点会重新执行一次簇头选举算法，

2 仿真结果分析

为了验证算法的有效性，利用 NS2 网络仿真工具^[8]对算法进行仿真，并结合 MATLAB 建立数学模型。仿真参数如表 1。其中 100 个节点随机分布在网络区域中，信道带宽为 1Mb/s，每个数据包为 500bytes，每个数据包头为 30bytes。

表 1 NS2 仿真环境主要参数

参数名称	参数值	参数名称	参数值
网络区域	100m×100m	d_0	87m
基站位置	(100,175)	数据包尺寸	4000bits
N	100	E_{elec}	50nJ/bit
初始能量	2J	ϵ_{fs}	10pJ/bit/m2
ϵ_{mp}	0.0013 pJ/bit/m4		

在文献[7]中，最优的簇头数目大约为 5。在形成簇阶段的距离门限值我们选为簇半径的 1.5 倍。这个门限值的选择不仅要考虑网络内簇的数目和簇的大小，还要保证能使簇头数目接近为 5，以及簇间区域不能相互重叠。网络每一轮循环时间为 60s。

我们从三个方面来比较 LEACH 协议与 CPRN 协议。

(1) 网络生命周期

我们通过第一个节点和最后一个节点的死亡时间间隔来评价三个协议的网络生命时间。图 1 显示相对于 LEACH 和 CPRN 协议，CPRN 通过较低的控制开销和有效负载均衡，有效的延长了网络生命时间。

(2) 基站接受到得数据量

和 LEACH 相比，CPRN 有效的均衡了网络的能量负载，降低了簇头的能量消耗，从而节省了网络的整体能量，节点存活时间长。基站因此可以接受更多的数据量。如图 2 所示。

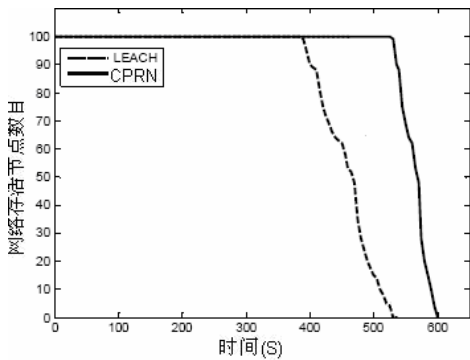


图 1 网络存活节点数

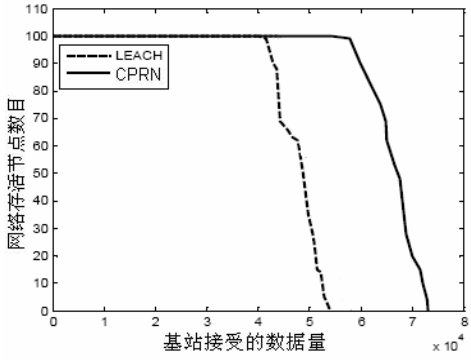


图 2 基站接受的数据量

(3) 网络中的簇头数量

在 CPRN 协议中对 LEACH 协议进行了改进，即当普通节点没有找到合适的簇加入时，它自己就当选为簇头，这样形成的簇头数目不会太小。相反，临近的竞争节点自动放弃竞争权利，必然导致簇头数目也不会太大。故由 CPRN 协议生成的簇分布合理，簇头数量少，簇与簇之间的覆盖区域交叠少。实验中我们用簇头之间的最小距离来做为簇头分布均衡的衡量标准。如图 3 所示。

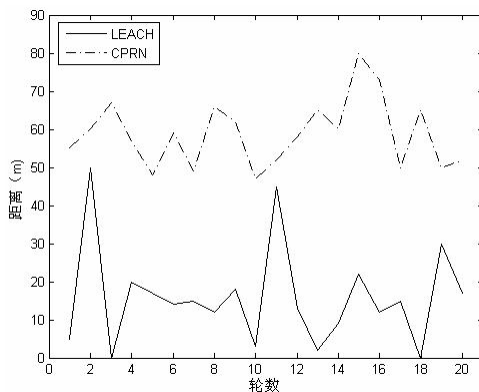


图3 网络中簇头分布

3 结束语

本文提出了一种协作传输的分簇路由协议 CPRN。在 CPRN 中，簇头均匀分布，花费的控制开销小。引入的协作传输机制有效的均衡了网络能量负载。同时使用的事件故障检测系统很好的提高了网络的鲁棒性。仿真结果显示，CPRN 协议能有效的降低网络的能量消耗，延长网络生命周期。

参 考 文 献

- [1] 孙利民,李建中,陈渝,朱红松,无线传感器网络[M].北京,清华大学出版社,2005.5
- [2] ESTRIN D.C and SRIVASTAVA M , “Overview of Sensor Networks”. IEEE Computer, vol. 37, no. 8, pp. 41-49, Aug. 2004.
- [3] FAN Yiming,YU Jianjun. The Communication Protocol for Wireless Sensor Network about LEACH[C]. Proceedings of 2007 International Conference on Computation Intelligence and Security Workshops. February 2007. 550 – 553.
- [4] BANDYOPADHYAY S,COYLEE J. An energy efficient hierarchical clustering algorithm for wireless sensor networks[C].Proceedings of INFOCOM 2003. San Francisco.2003.1713-1723.
- [5] 李成法,陈贵海一种基于非均匀分簇的无线传感器网络路由协议[J].计算机学报, 2007,30(1):88-91.
- [6] HEINZELMAN W, CHANDRAKASAN A and BALAKRISHNAN H. Energy- Efficient Communication Protocol for Wireless Microsensor Networks[C]. Proceedings of the 33rd International Conference on System Sciences (HICSS '00), Hawaii .January 2000.
- [7] HEINZELMAN W, CHANDRAKASAN A and BALAKRISHNAN H. An Application-Specific Protocol Architecture for Wireless Microsensor Networks, IEEE Transactions on Wireless Communications, vol.1, no.4 pp. 660-670, 2002.

作者简介:

李雷(1958-), 男, 安徽砀山人, 博士, 教授, 研究方向为智能信号处理、非线性分析与计算智能;

付东阳 (1983-), 女, 河南巩义人, 硕士研究生, 研究方向为无线传感器网络;

李艳辉 (1986-), 女, 河南鹤壁人, 硕士研究生, 研究方向为无线传感器网络。

DSTFT解调FSK信号中窗函数的影响

陈大平 陈忠辉

(福州大学物理与信息工程学院, 福建, 福州, 350002)

摘要: 提出了一种采用离散短时傅立叶变换 (DSTFT) 解调 2FSK 信号的方法, 分析了窗函数类型及窗口长度对解调性能的影响, 仿真结果表明, 采用合适的窗函数类型及合适的窗口长度能使解调性能有较大提高。

关键词: 离散短时傅立叶变换; FSK; 窗函数

Influence of Window Function in FSK Demodulation Based on DSTFT

Chen Da Ping Chen Zhong Hui

(College of Physics and Information Engineering, Fuzhou University, Fuzhou 350002, China)

Abstract: A BFSK demodulation method using Discrete Short-Time Fourier Transform (DSTFT) is presented. The demodulation performances impacted on by window function type and window length are analyzed. The test result indicates that using suitable window function and window length can greatly improve the demodulation performance.

Keywords: Discrete Short-Time Fourier Transform (DSTFT), FSK, Window function

1 引言

近年来, 时频分析这一新的信号处理方法因为其在分析非平稳信号时的优势而引起学者们的研究热潮, 而 DSTFT 做为时频分析中最简单, 最容易实现更为研究者研究的焦点。FSK 是一种常用的调制方式, 在工程中有广泛的应用。但在一些复杂的信道如电力线信道中, 由于噪声非常大, 用传统的 FSK 解调方法很难恢复出基带信号。传统的解调方法都是在时域进行的。而 DSTFT 解调 FSK 信号是在频域进行的。本文在前期研究在基础上, 分析了窗函数对解调性能的影响。通过仿真, 找到解调性能较好的窗函数及窗长度。

2 DSTFT解调 2FSK信号

2.1 离散短时傅立叶变换(DSTFT)

离散傅立叶变换(DFT)的定义为:

$$X(k)=\begin{cases} \sum_{n=0}^{N-1}x(n)W_N^{kn},0\leq k\leq N-1 \\ 0,其它 \end{cases} \tag{1}$$

离散傅立叶变换是对傅立叶变化的离散化。它不提供时间信息。因此， $X(k)$ 表示的是某一频率在整个信号持续时间内的 大小。对于非平稳信号来说，信号的每一次实现过程都不一样，仅仅知道信号频率的全局信息是远远不够的。

下面的两个数字信号 $x_1(n)$ 和 $x_2(n)$ 是两个不同的信号。

$$\begin{aligned} x_1(n) &= \frac{1}{2}\sin(10\pi n) + \frac{1}{2}\sin(20\pi n), 0 \leq n \leq N-1 \\ x_1(n) &= \frac{1}{2}\sin(10\pi n) + \frac{1}{2}\sin(20\pi n), 0 \leq n \leq N-1 \end{aligned} \tag{2}$$

$$x_2(n) = \sin(10\pi n)[u(n)-u(n-N/2)] + \sin(20\pi n)[u(n-N/2)-u(n-N)] \tag{3}$$

很明显，这是两个实现过程不一样的信号，那么它们的离散傅立叶变换分别如何呢？现在分别对这两个信号按相同采样率做相同点数的离散傅立叶变换，得到的结果如图 1 示：

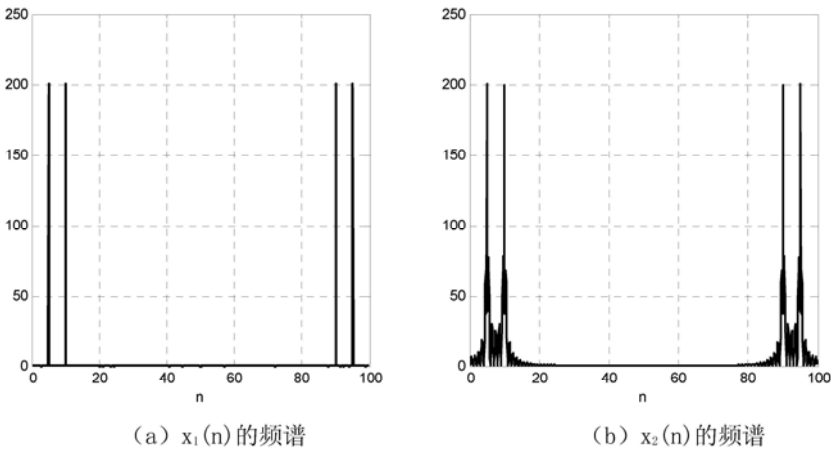


图 1 $x_1(n)$ 和 $x_2(n)$ 的频谱

从这两个图中很难分辨出这两个信号的区别，所以，傅立叶变换不适合用来分析非平稳信号。

离散短时傅立叶变换（DSTFT）的定义为：

$$X(n,k)=\sum_{m=0}^{N-1}x(m)w(n-m)\bullet\exp\{[-j2\pi(n-m)k]/N\} \tag{4}$$

其中， $x(n)$ 为输入信号序列， $w(n)$ 为移动窗函数， N 为窗宽度。从式（2）可以看出，DSTFT

是把信号在所关心的时刻附近取一小段，假设窗内信号是平稳的，对其做傅立叶变换，即可近似得到该时刻的瞬时频谱。当窗口移动后，又得出下一时刻的频谱。这样就把时间信息和频率时间信息联系了起来，得到了时间和频率的对应关系。DSTFT 与 DFT 最大的区别在于离散傅立叶变换不包含时间的信息，而 DSTFT 包含了时间信息，当 n 变换时，得到信号频率随 n 变化的规律。而 2FSK 信号的特点在于不同的信息用不同的频率来表示，因此，可以利用 DSTFT 来解调 2FSK 信号的。其工作流程如图 2 所示：

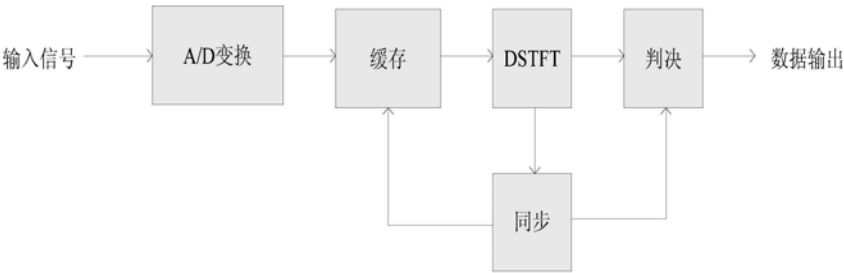


图 2 系统解调流程

2.2 判决算法

离散化的 2FSK 信号可以表示成：

$$\begin{aligned}
 x(n) = & A(i)[u(n-n_i)-u(n-n_i-N)]\bullet\cos(2\pi\frac{f_1}{f_s}+\varphi)+ \\
 & [1-A(i)][u(n-n_i)-u(n-n_i-N)]\bullet\cos(2\pi\frac{f_2}{f_s}+\varphi)
 \end{aligned}
 \tag{5}$$

其中， $A(i)$ 代表第 i 个码元， N 代表窗口宽度， f_1 和 f_2 是信号的两个载频（分别代表数据 0，1） f_s 是采样率。信号的 2 个载波频率点分别为 $K_1=[Nf_1/f_s]$ 和 $K_2=[Nf_2/f_s]$ （ $[\cdot]$ 表示取整）。频率点的幅度值为：

$$\left|X(nN,K_1)\right|=\left|\sum_{m=0}^{N-1}x(m)w(nN-m)\exp\{[-j2\pi(n-m)K_1]/N\}\right|
 \tag{6}$$

$$\left|X(nN,K_2)\right|=\left|\sum_{m=0}^{N-1}x(m)w(nN-m)\exp\{[-j2\pi(n-m)K_2]/N\}\right|
 \tag{7}$$

如果 $\left|X(nN,K_2)\right|>\left|X(nN,K_1)\right|$ ，判为码元 1，否则判为 0。

2.3 同步算法

DSTFT 解调 2FSK 信号时，同步最为关键。同步就是要使滑动窗的位置对准码元交替时刻。本文采用搜索最大峰值的同步方法。同步过程分为粗同步和细同步。粗同步时，每次计算 DSTFT 完窗口移动 $N_s/8$ （ N_s 为码元宽度），这样，一个码元宽度内最多移 8 次。记录下最大频谱峰值 $X_{\max1}$ ，并与设定的门限值比较，如果 $X_{\max1}<Q$ ，则判断不是有效信号，继续搜索信号；如果 $X_{\max1}\geq Q$ ，则认为接收到了 2FSK 信号，这时记录下最大频谱峰值对应的分析窗移动次数 m ，就实现了粗同步。

细同步是在粗同步的基础上，每次窗口移动一个点，搜索出 $|X(nN, K_2)|$ 或 $|X(nN, K_1)|$ 最大值对应的窗口位置，这样就实现了精确的同步。实现同步后，就把窗口移动步长变为一个码元宽度，窗口每移动一次，判决一次^[2, 3]，达到解调的目的。

3 窗函数的选择

3.1 不确定性原理

令 $x(t)$ 为有限长时变信号， $X(j\omega)$ 为其傅立叶变换， $x(t)$ 的有限时宽 Δt 和有限带宽 $\Delta\omega$ 定义为：

$$\Delta t^2 = \frac{\int_{-\infty}^{\infty} t^2 |x(t)|^2 dt}{\int_{-\infty}^{\infty} |x(t)|^2 dt} \tag{8}$$

$$\Delta\omega^2 = \frac{\int_{-\infty}^{\infty} \omega^2 |X(j\omega)|^2 d\omega}{\int_{-\infty}^{\infty} |X(j\omega)|^2 d\omega} \tag{9}$$

不确定性原理描述了信号的时宽和带宽之间的关系，即对于有限能量的任意信号其时宽和带宽的乘积总是满足下面的不等式：

$$\Delta\omega\Delta t \geq \frac{1}{2} \tag{10}$$

不确定性原理指出同时具有无限长时宽和带宽的信号是不存在的^[4]。

3.2 窗函数的选择

3.2.1 窗长度的选择

由不确定性原理看出，具有无限长时宽和带宽的窗函数是不存在的。窗口长度长，带宽就窄，反之，带宽就宽。因此，在选择窗函数的长度时，必须在时间分辨率和频率分辨率之间做一个折中。要取得好的频率分辨率，就必然要使信号的持续时间更长，即窗口长度变长。而对于 DSTFT 来说，窗口长度过长，就会使变化过快的信号失去时间信息。若窗口长度取得短，可以得到很好的时间信息，但是会使带宽变宽，频率分辨率上要付出代价。

3.2.2 窗函数类型的选择

文献[1]告诉我们，数据处理加窗时，主要产生了两个影响：一是加窗后的频谱是信号频谱的近似，二是产生了吉布斯效应。因此，在信号处理中，选取窗函数主要依据是：窗谱主瓣尽可能窄。二是尽量减少窗谱的最大旁瓣的相对幅度，也就是能量尽量集中于主瓣。不同的窗函数的时频特性有很大的不同。不同的窗函数具有不同的主瓣宽度和和旁瓣相对幅度，因此对解调性能有不同的影响。选窗口时，要求主瓣尽可能窄。同时又要求窗的最大旁瓣相

对幅度尽可能少，这样可以使能量集中于主瓣。然而，对任何一种窗函数来说，这两个要求是不可能同时满足的。主瓣宽度窄时，最大旁瓣相对幅度就大。当旁瓣相对幅度小时，主瓣宽度却会增加。因此实际中所选用的窗函数往往是他们的折中。常用的窗函数主要有：矩形窗、三角窗、汉宁窗、汉明窗和布拉克曼窗^[1]。表 1 给出了常用窗函数的指标。

表 1 常用窗函数的指标

窗函数	旁瓣峰值幅度/dB	主瓣宽度	阻带最小衰减/dB
矩形窗	-13	$4\pi/N$	-21
三角窗	-25	$8\pi/N$	-28
汉宁窗	-31	$8\pi/N$	-44
汉明窗	-41	$8\pi/N$	-53
不拉克曼窗	-57	$12\pi/N$	-74

从表 1 中可以看出，矩形窗具有最窄的主瓣宽度，但具有最大的旁瓣幅度。三角窗、汉宁窗、汉明窗都具有同样的主瓣宽度，旁瓣幅度也依次减小。不拉克曼窗具有最小的旁瓣幅度，但具有最大的主瓣宽度。上文说到，选去窗函数时，必须在主瓣宽度和旁瓣幅度间做折中，因此，可以预料选用汉宁窗和汉明窗的解调效果可能会比较好。

4 仿真结果

基于上述原理，在 matlab 上进行仿真。2FSK 信号的载波频率为 $f_1=102.4\text{kHz}$ ， $f_2=110.592\text{kHz}$ ，采样率为采样频率 $f_s=400\text{kHz}$ ，码元速率 $R=1000\text{ b/s}$ ，噪声为加性高斯白噪声。仿真比特数为 10^7 比特数。

窗函数分别采用矩形窗、挈比雪夫窗、汉宁窗、海明窗、布拉克曼窗、高斯窗。窗长度分别为 1.0、1.1、1.2、1.3、1.4、1.5 倍码元宽度。分别进行在相同窗口长度采用不同窗函数、同一种窗函数采用不同窗长度进行仿真，得出的结果如图 3、图 4 所示。

图 3 为当码元宽度为 1.3 倍码元宽度时不同窗函数仿真结果的误码率关系。

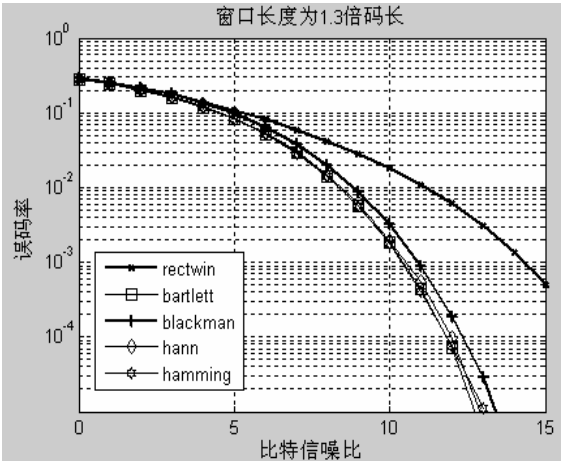


图 3 相同窗长度下不同窗函数的解调性能

从图 3 中可以看出, 矩形窗的解调性能与另外四种窗有明显的差别, 解调性能最差。这主要是因为矩形窗虽然有最窄的主瓣宽度, 但是它的旁瓣峰值幅度过大, 频谱泄露严重, 这严重影响解调性能。除去矩形窗外, 布拉克曼窗的解调性能略低于汉宁窗、三角窗、海明窗。汉宁窗、三角窗、海明窗的解调性能几乎是一样的, 难以分辨出其仿真性能的好坏。这也符合上文对窗函数特性的分析。

图 4 为采用汉明窗, 不同窗长度下的防真性能对比。可以看到, 窗长度过长, 窗长度过短, 都会使解调性能下降。当窗长度取在 1.3 倍码元宽度时, 解调性能比较好。

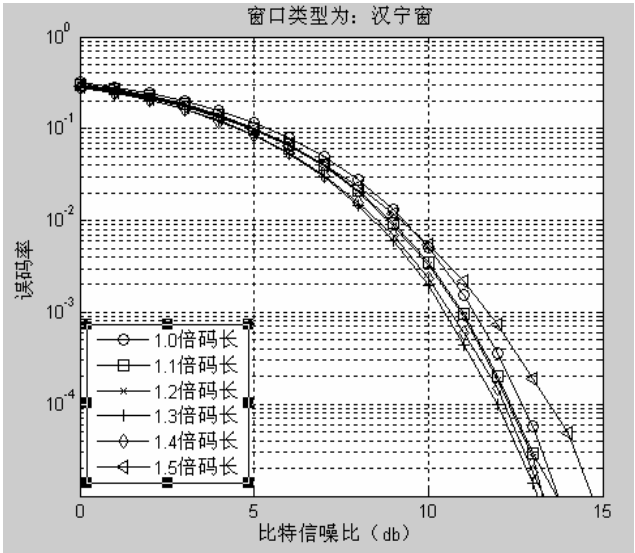


图 4 相同窗函数下不同窗长度的解调性能

5 结束语

本文分析了 DSTFT 解调 2FSK 信号时窗函数的影响, 从而找到一种最佳的窗函数以及最佳的窗口长度。通过仿真表明, 采用汉宁窗, 及窗口长度巍为 1.3 倍码元宽度时解调性能是最佳的。

参 考 文 献

[1] 刘顺兰, 等. 数字信号处理[M]. 西安, 西安电子科技大学出版社.
[2] 胡延平, 等. 一种基于 DSTFT 解调 2FSK 信号的新方法[J]. 通信学报, 2000, 21(6): 54~591.
[3] 林海斌, 等. 基于 DSTFT 的 FSK 信号的解调[J]. 现代电子技术, 2008, 31, 23.
[4] 张贤达 现代信号处理[M]. 北京: 清华大学出版社, 1996.

作者简介:

陈大平 (1983-), 男, 福建龙岩人, 硕士研究生, 通信与信息系统

陈忠辉 (1960-), 男, 福建福州人, 副教授, 通信与信息系统